Recognition of Chinese Medicinal Herbs By Gas Chromatgraphy

by

Suk Che Ho

A Thesis submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree of Master of Philosophy

in the

Division of Chemistry

of

The Chinese University of Hong Kong
(1998)

Thesis Committee: Prof. Alan Townshend (external examiner)
Prof. Suk Ping So (chairperson)
Prof. Ka Chung Hui (internal examiner)
Prof. Oi Wah Lau (supervisor)

Recognition of Chinese Medicinal Herbs by Gas Chromatography

Abstract

Chinese Medicinal Herbs have long been used for treating illnesses. It has been known that there are many kinds of herbs recorded in the Chinese Pharmacopoeia. Different kinds of herbs can have similar morphological features and smell that may be wrongly identified by these methods only. As they are widely used, it is important to identify correctly each kind of herbs for the proper use.

In this study, a recognition database system consisting of the library and searching methods has been developed, making use of gas chromatographic patterns of essential oils extracted from the herbal samples.

Since different herbal samples collected for this study can have great differences in volatile oil content, suitable "analysis window" was obtained through a "dilution strategy", by the proposed GC and dilution procedure. Analysis of essential oils was qualitatively and quantitatively studied by standardized GC/MS method. "Characteristic peaks" of herbal samples for database development were extracted from the GC retention data of "effective peaks", which in turn were chosen based on the proposed criteria.

In this project, 106 samples were examined by the proposed method. Promising results of recognition could be obtained. It was found that some herbal samples have similar searching results due to the close relationship between the species, such as same genus and similar therapeutic indices.

The approach introduced here offers a simple, user-friendly and systematic way for the recognition of Chinese Medicinal Herbs containing essential oils. It allows us to recognize the herbal sample, requiring no special professional knowledge in Chinese Medicinal Herbs. The MS data obtained from GC/MS analysis can provide a valuable supportive information for the future work of using direct MS patterns for recognition.

撮要

使用中草藥治療疾病已有著相當悠久的歷史。在中國藥典裡，輯錄了很多不同類別的中草藥，每一種均有其特徵，也有些有著相似的外形及味道，若單靠它們的外表特徵來作分辨，有可能會出現錯誤的情況。基於中草藥被廣泛使用，正確辨認它們是相當重要的。

這個研究是建立一個辨認中草藥的方法。這方法是選用草藥裡揮發油的氣相色譜圖，加以分析，得到數據，來建立資料庫及辨認的系統。

不同的草藥樣本的揮發油含量差別可以很大，通過氣相色譜的分析步驟及稀析程序，可得到適合的「分析視窗」，以用作可行的數據比較。通過統一的氣相質譜方法，可定性及定量地分析揮發油的成份色譜圖。由此，利用研究中所定立的條件，一些「有效」色譜峰，可以被抽取出來，繼而進行另一輪「特性」色譜峰的抽取，而色譜峰的相對保留時間及歸一面積亦可透過運算式計算出來。通過所定立的相配方法及計分方法，樣本便能被辨別出來。

在此項研究中，進行了一百零六個樣本分析，從數據看來，發覺所建立的方法可取得頗佳的辨認結果。此外，一些草藥樣本因有相近的關係（如同屬及具相似療效），它們的色譜圖會較爲相似，因而得到近似的辨認結果。

這裡介紹了一個簡單而有系統的方法來辨認含有揮發油的中草藥，並不需要對中草藥有專業的知識，而由氣相質譜分析得來的質譜數據也能提供有用的資料，以助將來用直接質譜分析方法來辨認中草藥。

Acknowledgments

I would like to thank Dr. Oi Wah Lau, my project supervisor, for her sincere encouragement and kind concern, not only on my project, but also in other aspects throughout this study. My sincere thanks also goes to Dr. Tao Bai for his valuable suggestions and advice. His kind care and support not only encourage me to face difficulties in my research, but also give me confidence to try my best during the study.

I would like to thank the following people: Dr. Bing Shao for his sincere advice, care and encouragement during my research study; Ms. Karen Wai Yin Fong for her conceptual assistance, sincere help and care, and for her patience though I bothered her a lot; Mr. Kwok Kwong Ng for his technical support during the course of this research; my dear brothers and sisters in the Post Graduate Ministry of the Campus Crusade for Christ, for their sincere care, encouragement and prayers; Mr. Joseph Chun Kong Chan for his invaluable care and support, and for his patience to listen to my complaints and always walk beside me; and Ms. Chris Yee Ting Kwan, Mr. Ho Yin Leung, Ms. Michelle Tin Wun Lee and Ms. Avis Chun Yin Mak, colleagues and friends, for their enthusiastic support.

Finally, I would like to thank my Lord for HIS love and promises:
"My grace is sufficient for you, for my power is made perfect in weakness."

2 Corinthians 12:9

Dedication

To

My Parents

Abbreviations

CMH	Chinese Medicinal Herbs
GC	Gas Chromatography
MS	Mass Spectrometry
GC/MS	Gas Chromatography/Mass Spectrometry
FID	Flame Ionization Detector
ITD	Ion Trap Detector
psi	Pounds Per Square Inch
ppm	Parts Per Million
IS	Internal Standard
SD	Standard Deviation
w.r.t.	With Respect To

Table of Contents

Chapter 1: Introduction

1.1 Overview of Chinese Medicinal Herbs containing essential oils

1.1.1 Introduction of Chinese Medicinal Herbs

Chinese Medicinal Herbs (CMH) have long been used for treating illnesses. In traditional Chinese herbology, the term herbal medicine includes mineral and animal products as well as plants, with plants constituting the majority. In the present study, the selection was limited to medicinal plants only. Chinese Medicinal Herbs can be classified by different focal points. Four kinds of classifications are adopted, with focal points on (a) therapeutic function, (b) herbal parts with pharmaceutical value, (c) bioactive components, and (d) natural relationship such as genes [1].

It has been known that there are many kinds of herbs documented systematically in Chinese Pharmacopoeia, with the information of their physical features, chemistry, therapeutic indices, etc. CMH have been used in cardiovascular, nervous, alimentary, respiratory and hematopoietic systems and chemotherapy. Up till now, still over half of the Chinese population relies on herbal prescriptions rather than Western Medicine, particularly when Western medicines do not produce the desired results [2]. In the past 40 years, a great deal of work in the compilation of Chinese drugs, recipes and production technology of crude drugs and traditional patent medicines has been done [3]. Owing to the rapid utilization of scientific methods in this area, increasing interest has been shown in studying Chinese medicines. Advances in instrumentation and analytical methods permit more information to be discovered with the use of only small amounts of samples.

1.1.2 Chinese Medicinal Herbs containing essential oils

Chinese Medicinal Herbs are very complex and they generally contain proteins, sugars, glycosides, essential oils, etc [1, p.80].

Essential oils of a large number of plants exhibit useful biological, pharmacological and therapeutic activities and the oils can be extensively utilized in the preparation of prescription and non prescription drugs. In addition to their pharmaceutical uses, essential oils are commercially important chemical compounds. They are generally complex mixtures of naturally occurring compounds or secondary metabolites and their utilization in the various industries is influenced by the nature of their constituents [4, p.69].

Essential oils from CMH are mixtures containing volatile components which can be steam-distilled from the herb tissues. Generally, the essential oils are in liquid form and can vaporize at room temperature. The oils can dissolve in organic solvents such as ether easily but not dissolve in water [5]. The chemical composition of essential oils are typically dominated by mono- and sesquiterpenoids with minor amounts of phenyl propanoids, diterpenoids and various organic compounds, which are steam volatile [4]. Most of them have fragrant smell and odoriferous. Their therapeutic values have been known since ancient times [6, p.159]. Some ingredients of essential oils, such as camphor, are found to be pharmaceutically active. Thus, essential oil from CMH is a valuable source of potential pharmaceuticals, which is worthwhile to be investigated.

1.2 Recognition of Chinese Medicinal Herbs

There are over 2,000 items listed in Chinese herbal pharmacopoeias, but only about 300 are used in general practice, of which less than one hundred are regarded as indispensable in formulating the most popular prescriptions [7, p.7]. Different kinds of herbs have their characteristic features such as morphological features, smell

and odor. However, some herbal medicines can have similar appearance such as Chuanwu (川烏) and Chuanxiong (川芎) that may cause confusion by just comparing their appearance without professional knowledge. Herbal medicines with similar appearance may have totally different therapeutic indices or even worse, some herbs are toxic that may cause death if a large dosage is taken. One such example was the incident concerning Weilingxin (威灵仙) and Guijiu (鬼臼) which happened a few years ago. Another problem in studying Chinese herbs is that numerous herbs bearing the same name are actually different. Conversely, one herb may frequently be known by several names [2]. Furthermore, after treatment of the herbal medicines such as grinding, the entire plant for macroscopic examination is usually not available. Nevertheless, due to the widespread use, it is important to recognize each kind of herbs for the proper use.

1.2.1 Traditional method in recognition of Chinese Medicinal Herbs (CMH)

Traditionally, for the recognition of CMH, macroscopic and microscopic descriptions as well as physiochemical tests for identity are carried out. Macroscopic descriptions include shape, size, color, surface and fracture characteristics, smell and taste [3, p.233; 8, p.18]. The disadvantages of these methods are that skill and experiences are needed, and that no regular patterns can be followed [9, p.4]. The method of taxonomy can also be used for recognition, however, it also requires experience and professional knowledge. Moreover, the two aforementioned methods are limited to the availability of the entire plants. Microscopic descriptions include histological structures, cell types and cell contents, which are only visible with the aid of a microscope [3, p.234]. It was found that micro-analysis also have some limitations that it may be difficult or impossible to distinguish plants which are near relatives due to the very similar cell structures [9]. Training and rich experiences are also needed. Physicochemical tests include simple chemical qualitative reactions, examination under ultraviolet light, microsublimation, and various chromatographic methods [3, p.234].

1.2.2 Instrumental methods for the recognition of CMH

For identification of CMD, thin layer chromatography (TLC) has been widely used, where the comparison of the retention properties of the unknown are compared with those of the reference samples. However, its greatest disadvantage is that the reproducibility of the R_f values is not good. Another problem is that because a lot of developing agents are used for different samples tested, it is difficult to standardize the method.

Recognition of CMH based on ultra-violet spectra has been proposed [10]. It has been pointed out that the UV spectrum of a species of CMH is relatively stable and reproducible. However, most herbal drugs only show few absorption peaks in their spectra, the information used for recognition is poor and it was reported that only 32% of the tested herbal drugs have two or more peaks. Derivative ultra-violet spectra were also used for the purpose of recognizing CMH [11]. It was shown that the herbal drugs of different families and genera have specific spectral features in the derivative spectra. The peak values of the derivative spectra are the reliable information of recognition.

The specific character of IR spectra for recognizing herbal drugs is superior to that of the UV method. IR spectrometry has been used for the recognition of some CMH. The experiment demonstrated that if a fixed extraction method is used, the IR spectra of the extracts are reproducible [12].

1.2.3 The use of GC and GC/MS on CMH

Gas chromatography (GC) represents a useful and fast method of separating the individual components of a mixture. It is the combination of rapidity and resolving power which makes GC superior to all other chromatographic methods. For instance, when comparing thin layer chromatography (TLC) with GC, the

resolution of the various constituents of an essential oil by TLC remains less than satisfactory.

Because of the complexity of the samples, high-resolution gas chromatography (GC) is the most attractive method for the determination of trace organic compounds with sufficient volatility. Detailed information about both the qualitative and quantitative composition can be obtained within a reasonable time; qualitative analysis, if required, can be effected in combination with mass spectrometry (MS) [8]. GC and MS are promising analytical tools. Analytical chemists can detect minor and trace components in the nanogram range [13, p.156].

For CMH, GC is often used in several aspects. For examples, it helps to study the chemical composition, qualitatively and quantitatively, of the essential oils extracted from the herbal drugs. Besides, it is surely one of the most efficient tools for quality control of the materials and so, checking the purity or to overcome cases of fraud is possible. In addition, it helps to distinguish features which aid in the identification of the various hybrids. Furthermore, GC can help to detect impact chemicals which are responsible for the therapeutic function. With the combination of MS, the bioactive components can be structuralized that novel drugs can be investigated to treat illness.

1.3 Motivation and objective of this research

As advances in instrumentation and analytical methods now permit spectral analyses of very small quantities, a simple recognition methodology has been developed in this research for assessing the possibility of recognizing Chinese Medicinal Herbs with scientific methods. To develop the method, gas chromatographic patterns of essential oils extracted from herbal samples were under study.

1.3.1 Motivation

On the review of previous work, it was found that GC/MS usually applied on the study of the identities and abundance of the components inside the essential oils of herbal drugs. Besides, studies of the chemotypes among the species with close relationships have been carried out. In this area, statistical methods have often been used to differentiate species with close relationship or the same species grown in different places, time or conditions. These were appreciable work as differentiation of similar chromatographic patterns could be carried out. However, prior to this kind of differentiation, a scientific and systematic scheme for the species recognition between broader and various kinds (between genera and among genus) of herbs has not been studied. Actually, this kind of recognition is valuable because it can serve as a pre-recognition or a pre-filter process before further differentiation. The science-based method can provide a standardized procedure that rich experience and professional knowledge on CMH is not necessary. Moreover, the relationships between different genera or species can be drawn in a broader sense.

In recent studies, it was found that much effort have been put on chromatographic pattern recognition for the analysis of complex mixtures. Essential oils from flowers and plants are also a kind of mixture under investigation.

As previously mentioned, essential oils are important raw materials for a number of industries and their use is widespread and varied. Quality control as well as research into essential oils was based on organoleptic evaluation by trained test panels combined with analytical data on major as well as on minor compounds and even on traces, obtained with modern techniques. The advent of gas chromatography has considerably accelerated the study of the composition of essential oils [13].

Modern capillary gas chromatography shows powerful resolution even for complex mixtures. In particular, components of essential oil can be effectively separated on a certain stationary phase in a capillary column, showing patterns

consisting of well resolved peaks. Indeed, there is much valuable information concealed under the complicated GC patterns [14], which are specific as they can be used as a fingerprint of certain crude drugs containing essential oils. For example, the principal Cinnamomum species of commerce can be identified in powdered samples based on differences in their characteristic chromatographic profiles [20]. It is possible to distinguish herbal samples from different origins. The recognition of these patterns on the chromatogram can be accomplished by comparing the chromatograms of an unknown complex mixture with that of a standard essential oil with the use of a data processing system designed with a personal computer program [15].

Furthermore, with the combination of MS, GC/MS can help to give more information from the chromatographic patterns. The identification of individual component is possible by comparing its retention index and mass spectrum with those of the corresponding reference chemical. The relative abundance of the components can also be estimated. Besides, it can provide mass information of the composition of extracted essential oil.

1.3.2 Objective of this research

The reliable identification of a complex mixture is always a tedious task for the analyst. In this research, the merit of the possibility in recognizing chromatographic patterns of essential oils was applied in the field of Chinese Medicinal Herbs, in which a systematic scheme for the species recognition based on scientific method has been developed. In order to make recognition possible, standardization of the method was required before comparison can be made. Thus, in the study, the experimental procedure including pretreatment of the samples, essential oil extraction, instrumental parameters and settings, and data analysis methods were standardized. A recognition database system consisting of the library and searching methods has been developed, using chromatographic patterns of essential oils obtained by GC/MS.

The methodology has been built up to assess the possibility of recognizing Chinese Medicinal Herbs (which contain essential oils) in a broad sense, i.e. between genera and species. Further or detailed differentiation of the subspecies, variates or herbs with very close relationships, were not under investigation. The identities of the herbal drugs recognized were given in the form of common or commercial names, unless accurate scientific names were known.

The methodology tends to offer a simple and systematic way for recognition requiring no special professional knowledge on CMH. The method can be applied to various kinds of herbal drugs, provided that they contain essential oils. Small amounts of samples can be used. Sometimes, the variation in the contents of essential oils among various kinds of herbs can be so large that overloading problems may occur when using GC/MS. In this methodology, a dilution strategy is introduced to provide comparable "analysis window" to overcome this problem. This allows the utilization of the methodology in both herbal samples with high as well as low oil contents. Furthermore, through analysis of the essential oils by GC/MS, the GC retention indices and the MS data can provide information on the composition of essential oils apart from the recognition of the identity of the herbal drugs. With this information, relationships or therapeutic values among herbal drugs may be studied.

1.4 Outline of the methodology and arrangement of the thesis

The overview of the method is shown in Figure 1.1.

In order to make comparison possible, the procedure has to be standardized. Samples were preliminarily treated to obtain a suitable size for extraction. Essential oils from herbal samples were extracted using a Dean and Stark-type trap apparatus. *n*-Hexane was used as solvent to concentrate the essential oils in aqueous condensate [15]. Pretreatment of the herbal samples and extraction of the essential oils will be discussed in greater details in Chapter 2.

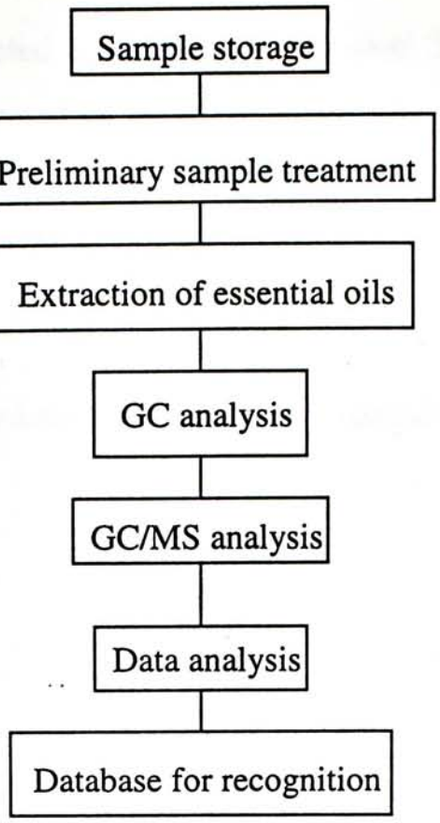


Figure 1.1 Overview of the method

After dilution and addition of internal standards, the diluted essential oils were analyzed by gas chromatography. GC could provide information about the number and abundance of the components inside the essential oil. More importantly, it helped to estimate the oil abundance in the herbal sample, from which the most suitable dilution strategy for the extract would be adopted to provide suitable "window of analysis" for subsequent GC/MS analysis. GC/MS analysis could provide information of GC retention indices which served as codes and were used to build up the database. Besides, MS data served as an auxiliary information for choosing peaks from the chromatograms. Details of the instrumental analysis of the essential oils will be discussed in Chapter 3.

Chapter 4 is concerned with the database development for recognition, which consists of the library and searching systems. In the library system, the retention indices and the abundance of the peaks were analyzed. "Characteristic" peaks of

herbal samples were extracted from the GC retention data of "effective" peaks, which in turn were chosen based on the proposed criteria. In the searching system, a matching method was developed to give a similarity score between the sample and the target candidate. The performance of the recognition method is discussed in Chapter 5.

Finally, some conclusions about the overall approach of the methodology will be described in Chapter 6.

Chapter 2: Experimental Setup

2.1 Reagents and materials

2.1.1 Reagents and glassware

The reference chemicals were purchased from Aldrich, Merck and Sigma Chemical Companies with purity of about 95% or above. Ethyl caprylate and ethyl caprate [16,17], with >99% purity, were used as the internal standards. HPLC grade *n*-hexane was used in extraction and was tested for any impurities present before use. Ultra-pure deionized water (with filtered and low pressure reverse osmosis treatment) was used in extraction and washing. Boiling chips used in extraction were washed with ultra-pure deionized water, HPLC grade acetone and *n*-hexane, and dried at 100°C before use. Analytical grades sodium chloride and anhydrous sodium sulfate were used after extraction. All glassware was washed thoroughly with detergent and rinsed with ultra-pure deionized water and HPLC grade acetone. They were dried and rinsed with HPLC grade *n*-hexane before use in order to minimize contamination. The dilution factor for each dilution was not more than 100 fold.

2.1.2 Materials

In the study, 37 species of herbal medicines (106 herbal samples in total) containing essential oils were analyzed. The herbal drugs under analysis were between genera and among genus. Some of them have similar therapeutic functions. Since the accurate scientific names of some retail samples were not known, those with the same retail names were grouped as one kind of species. Samples from different sources were investigated. It was assumed that the sources of the herbal samples purchased from different retail stores were different. Since herbal samples from different botanical and geographical origin can affect the relative abundance of the components in the oil, it is important to study different sources in order to

"extract" any "characteristic" peaks from the chromatographic patterns for a particular herb.

The 37 species of herbal medicines were listed in Table 2.1. Some samples were purchased from Chinese medicines retail stores in Guangzhou, Beijing and Mongolia. Others were purchased from stores in Hong Kong. Among the 106 herbal samples, 16 were purchased from a laboratory in China selling authentic herbal drugs. These samples were provided with accurate scientific names.

Table 2.1. Herbal medicines studied

Common names used in Chinese Pharmacopoeia (code)*	Retail names or scientific names	Number of sources analyzed
Danggui (01) 當歸	Danggui	4
	Angelica sinensis (Oliv.) Diels	1
Duhuo (02) 獨活	Duhuo	6
	Angelica pubescens Maxim. F. biserrata Shan et Yuan	1
Qianghuo (04) 羌活	Qianghuo	3
	Notopterygium incisum Ting ex H. T. Chang	2
Baizhu (05) 白朮	Baizhu	4
	Atratylodes macrocephala Koidz	1
Canzhu (06) 蒼朮	Canzhu	4
	Atratylodes lancea (Thumb.) DC.	1
	Atratylodes chinensis (DC.) Koidz	1
Jingjie (07) 荆芥	Jingjie	8
	Schizonepeta tenuifolia Briq.	1
Bajiaohuixiang (08) 八角茴香	Bajiao	1
	Illicium verum Hook.f.	1
Sharen (09) 砂仁	Sharen	4
	Amomum villosum Lour.	1

(cont'd)

Ezhu (11) 莪术	Ezhu	4
	Curcuma phaeoculis Valeton	1
	Wenezhu (Curcuma wenyujin)	1
Yujin (12) 郁金	Yujin	4
	Wenyujin (Curcuma wenyujin)	1
	Pianjianghuang (Curcuma wenyujin Y. H. Chen et Ling) (Rhizoma Wenyujin concisa)	2
Chuanxiong (15) 川芎	Chuanxiong	6
	Ligusticum chuanxiong Hort.	1
Qianhu (18) 前胡	Qianhu	3
	Peucedanum praeruptorum Dunn	1
Fangfeng (19) 防風	Fangfeng	4
Muxiang (20) 木香	Muxiang	8
	Aucklandia lappa Decne.	1
Zisuye (23) 紫蘇葉	Zisuye (purplish green)	6
	Zisuye (greenish)	3
	Perilla frutescens (L.) Britt.	1
Xiangru (24) 香薷	Xiangru	3
Gaoliangjiang (25) 高良姜	Gaoliangjiang	4
Peilan (29) 佩蘭	Peilan	4
Huoxiang (30) 霍香	Huoxiang	4
	Total species: 37	Total samples: 106

* Codes are designated for each herbal drug for database development

2.2 Sample pretreatment

Samples were crushed by using a stainless steel blender to obtain coarse crushed samples [18] which were then crushed by another blender for further grinding of the solids into finer powder. Suitably sized samples were obtained by passing the powder through a 30-mesh sieve for extraction. The powder sample was then stored in a clean dry plastic bottle ready for use.

2.3 Extraction of essential oils from the herbal samples

2.3.1 Traditional extraction methods for essential oils

The most commonly used techniques for extraction of essential oils from herbal material are steam or hydrodistillation and solvent extraction. The essential or volatile oils that are commercially available are generally obtained using steam distillation. Other methods include headspace analysis techniques [19, p.142].

For steam distillation, two kinds of technique are usually carried out in two slightly different ways. In the first, samples are covered with water in a suitable vessel fitted with a condenser, and while the mixture is gently boiled, a condensate is collected (hydrodistillation) and the oil then separated from the water. This kind of technique is adopted in Chinese Pharmacopoeia. In the second, steam is passed through a vessel containing a sample-water mixture (steam distillation) to yield a similar condensate. For all classes of experiment carried out under the heading of steam distillation, the plant material is subjected to temperatures up to 100°C [5, 19].

For solvent extraction, the plant material is soaked in organic solvents at ambient temperatures or below, or extracted by a boiling solvent in a Soxhlet-type apparatus. Recovering the volatile oils requires evaporation of the solvent, and care must be taken that the more volatile compounds are not lost. In this method,

materials of low volatility are also extracted out that further distillation is required [5, 19].

Solvent extraction techniques are efficient for isolating the essential oil fraction from herbal species but provide an extract that is contaminated with materials of low volatility unsuitable for separation by high resolution gas chromatography without further sample preparation steps. Steam distillation and headspace analysis techniques are alternative approaches uniquely suited to separations by gas chromatography because they provide an isolate that has a limited volatility range and would not normally contaminate the chromatographic system [20].

2.3.2 Extraction by hydrodistillation using Dean and Stark-type trap apparatus

In this study, essential oils were obtained by hydrodistillation using a Dean and Stark-type trap apparatus. Samples of crushed material were first wetted and swirled (4 g of material and 200 ml ultra-pure deionized water) in a round-bottomed flask for an hour. The wetted samples were then hydrodistilled in a Dean and Stark-type trap apparatus for 2 hours with washed boiling chips. The *n*-hexane (2.00 ml with HPLC grade) inside the trap was used to concentrate the essential oils in the aqueous condensate [15]. Condenser circulated with ice water was used in order to minimize the vaporization of the volatile components out of the apparatus. The setup of the apparatus is shown in Figure 2.1.

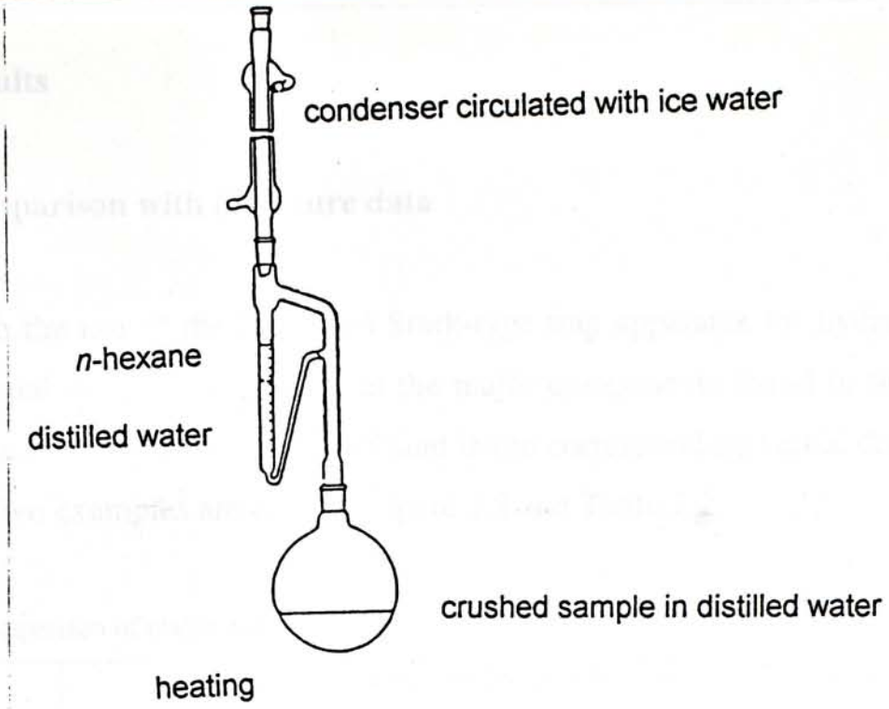


Fig.2.1. Dean and Stark-type trap apparatus used for hydrodistillation of herbal samples

The main advantage of this setup was that no further enrichment by solvent evaporation was required. Together with the extracted components, the solvent layer was drawn out, salted out by sodium chloride (AR grade) [5] and dried with anhydrous sodium sulfate (AR grade). The extract was placed in a vial sealed with paraffin film and stored at 4°C [21]. It was ready to be analyzed.

To analyze the oil, 0.10 ml of the crude extract was diluted by 10 fold with *n*-hexane and 0.10 ml of two internal standards (ethyl caprylate and ethyl caprate) from stock solution (450 ppm) were added in the diluted extract with final concentrations of 45 ppm. Two internal standards were added to enable the calculation of the relative retention indices of chromatographic peaks instead of the absolute retention times (see section 4.2.2). In the remaining part of the thesis, calculation related to abundance was referred to the peak of internal standard ethyl caprylate. Together with ethyl caprate, they were used to calculate the relative retention indices, but ethyl caprate was not involved in the calculation of quantitative information. The diluted sample was analyzed by gas chromatography.

2.4 Results

2.4.1 Comparison with literature data

With the use of the Dean and Stark-type trap apparatus for hydrodistillation of the essential oils, it was found that the major components found in the extracted oil were usually comparable to those found in the corresponding herbal drug from the literature. Two examples are cited in Figure 2.2 and Table 2.2.

Table 2.2. Comparison of major components between studied samples and literature

Studied sample	Some known components with high abundance	
	In extracted oil by proposed method	Found from literature information
Sharen	Bornyl acetate, camphor, borneol	Bornyl acetate, camphor, borneol [22]
Baizhu	Atractylon	Atractylon, γ-elemene, γ-cadinene, γ-patchoulene [23]

2.4.2 Reproducibility of the extraction

For the proposed extraction method, it was important to study its reproducibility, recovery, simplicity and economic aspect, etc. Among these factors, reproducibility of the method was the most important factor in order to allow efficient comparison and make pattern recognition possible.

In this study, the extraction process of both the mixtures of reference chemicals and herbal samples were carried out. The reproducibility of the extraction of reference chemicals was monitored by GC while that of herbal samples were by GC/MS. In the calculation of reproducibility, variations due to the preparation of the sample, the extraction process and the GC analysis were included.

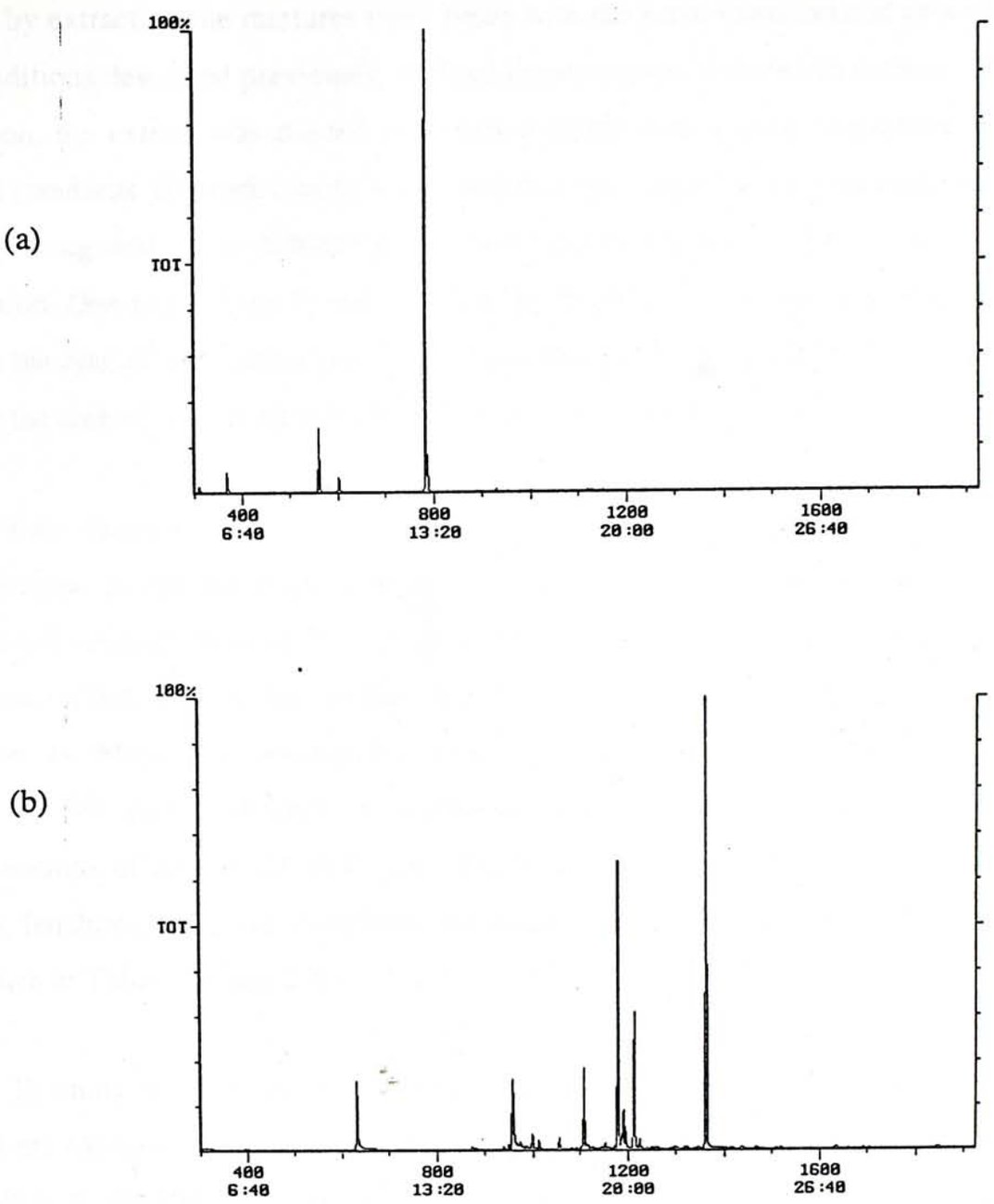


Figure 2.2. (a) GC/MS chromatogram of Sharen (b)GC/MS chromatogram of Baizhu

For the artificially prepared mixtures, reproducibility of extraction was studied by extracting the mixtures three times with the same experimental procedure and conditions described previously, without consideration of the matrix effect. After extraction, the extract was diluted to a certain extent with known concentration of internal standards. Reproducibility was calculated by comparing the area ratios of the corresponding analytes in different trials. Two types of information were used for the calculation. One type (Type 1) was to calculate the ratio of the area of each analyte peak to the area of the internal standard. Another type (Type 2) was to calculate the ratio of the area of each analyte peak to the total area obtained from all analyte peaks.

Two mixtures of reference chemicals were prepared. One with lower concentration of the reference chemicals to mimic the herbal samples with low essential oil contents (Mixture 1). Another with higher concentration of the reference chemicals which mimic the herbal samples with higher essential oil contents (Mixture 2). Mixture 1 contained 7 reference chemicals with concentrations of around 1,000 ppm. Mixture 2 contained the 7 reference chemicals with concentrations of around 120,000 ppm. The 7 chemicals were α-pinene, myrcene, cineole, fenchone, trans-caryophyllene, humullene and cinnamyl acetate. The results are shown in Tables 2.3 and 2.4.

To study the reproducibility of extraction of the herbal samples, three kinds of Chinese Medicinal Herbs were investigated: herbal sample with high oil contents (with $R > 5$, for the information of R, see section 3.1.3.3), herbal sample with moderate oil content and herbal sample with low oil contents(with $R < 1$). They were Jingjie, Zisuye and Duhuo respectively. The extraction of each herbal sample was done in triplicate. The corresponding chromatograms are shown in Figures 2.3 to 2.5.

Table 2.3. Reproducibility of extraction of mixture 1

Reference chemical	Type 1 information		Type 2 information	
	Ave. rel. area (%) of 3 extr.	Deviation at 90% CL	Ave. rel. area (%) of 3 extr.	Deviation at 90% CL
α-pinene	143.7	0.022	20.6	0.013
Myrcene	84.1	0.039	12.0	0.011
Cineole	120.0	0.109	17.1	0.011
Fenchone	110.5	0.154	15.8	0.016
Trans-caryophyllene	96.4	0.203	13.8	0.026
Humullene	94.5	0.197	13.5	0.026
Cinnamyl acetate	50.3	0.180	7.2	0.023
Average	99.9	0.129	14.3	0.018

Note: The average relative deviations at 90% confidence level of type 1 and type 2 information were 12.9% and 12.6% respectively.

Table 2.4. Reproducibility of mixture 2

Reference chemical	Type 1 information		Type 2 information	
	Ave. rel. area (%) of 3 extr.	Deviation at 90% CL	Ave. rel. area (%) of 3 extr.	Deviation at 90% CL
α-pinene	1152.9	2.160	15.0	0.017
Myrcene	665.2	0.769	8.6	0.002
Cineole	1231.2	1.264	16.0	0.005
Fenchone	1217.7	1.165	15.8	0.004
Trans-caryophyllene	1297.9	1.192	16.9	0.002
Humullene	1285.5	1.132	16.7	0.004
Cinnamyl acetate	838.0	0.599	10.9	0.011
Average	1098.3	1.183	14.3	0.006

Note: The average relative deviations at 90% confidence level of type 1 and type 2 information were 10.8% and 4.2% respectively.

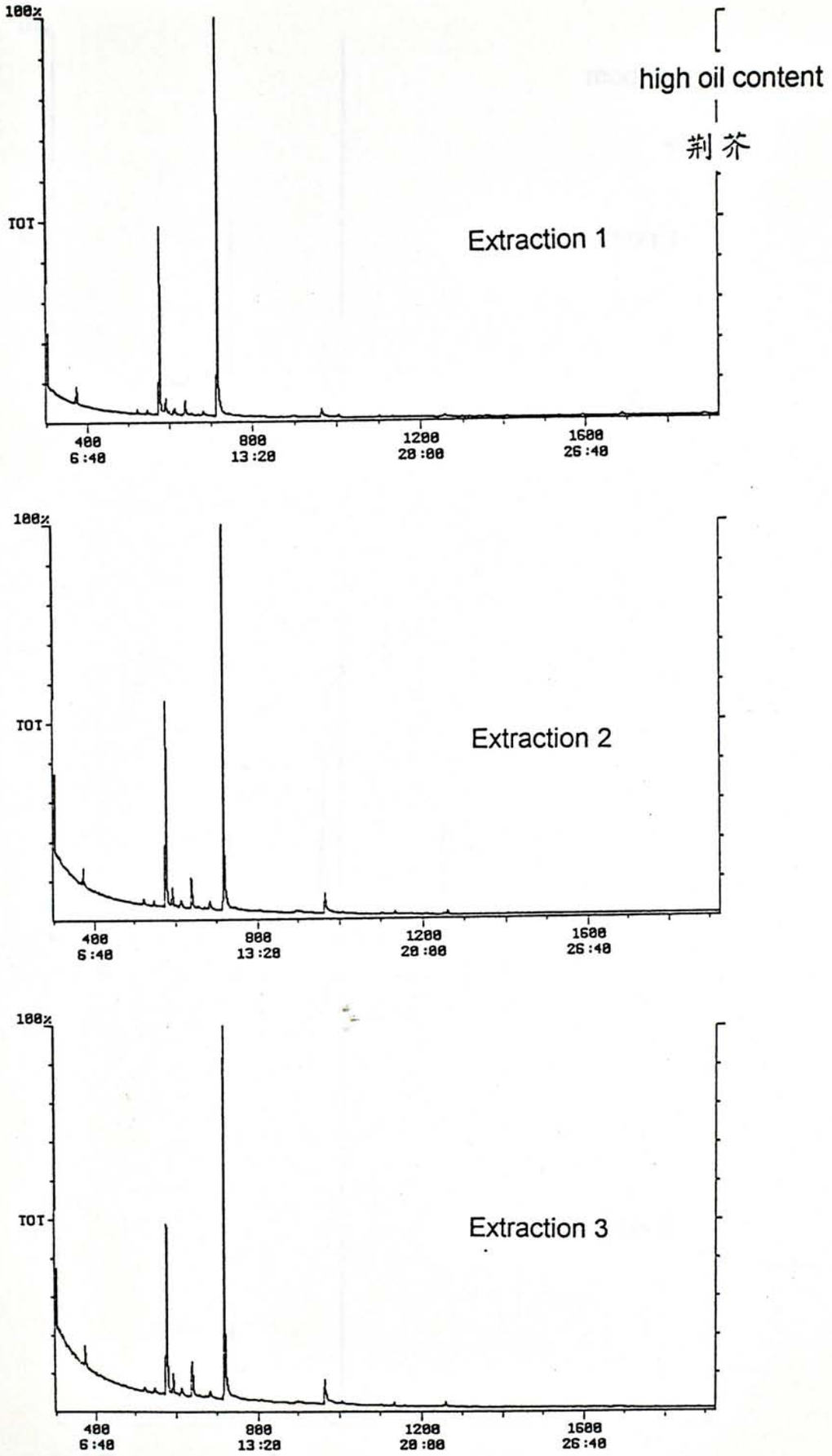


Figure 2.3. Reproducibility of extraction for Jingjie

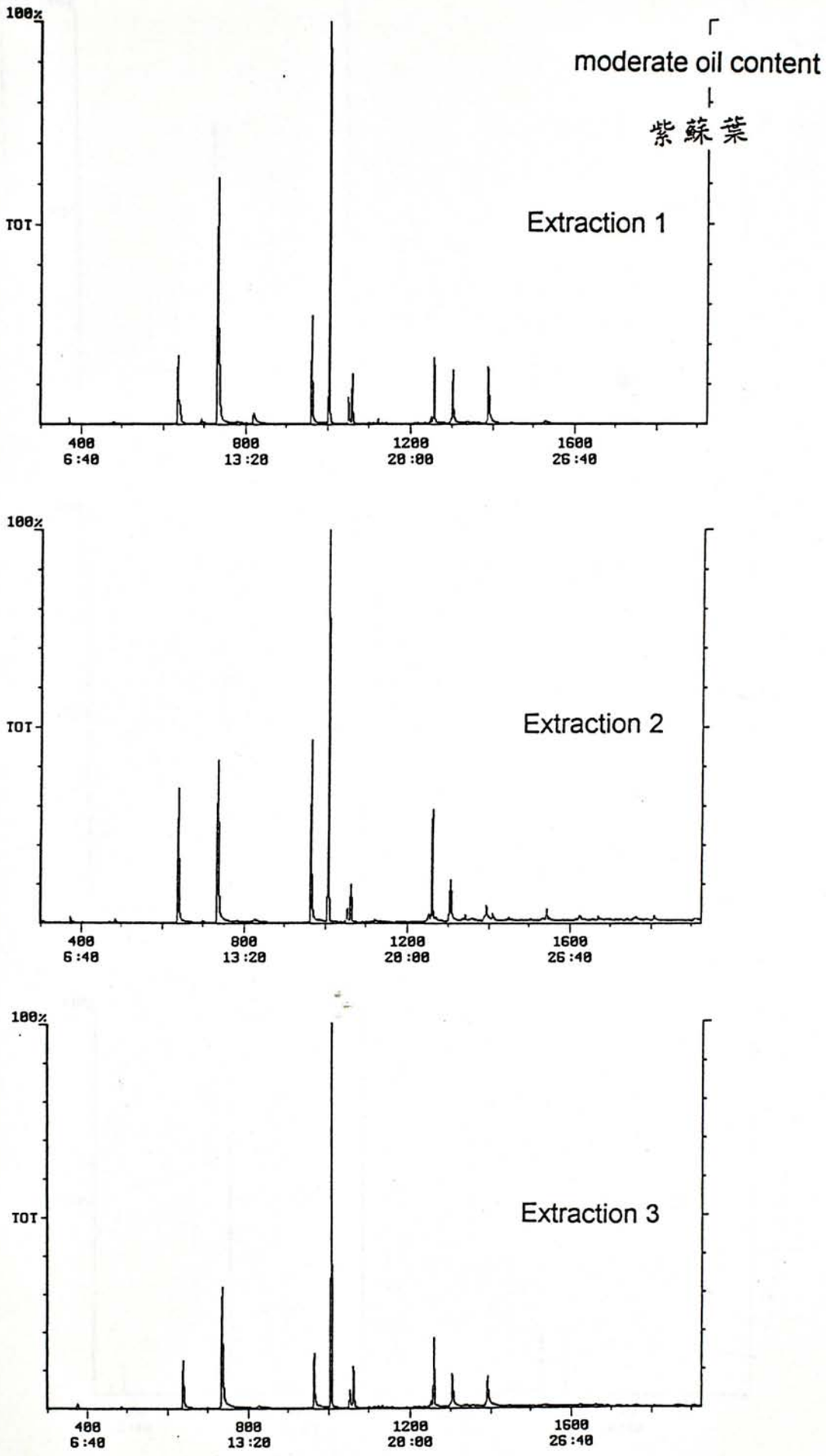


Figure 2.4. Reproducibility of extraction for Zisuye

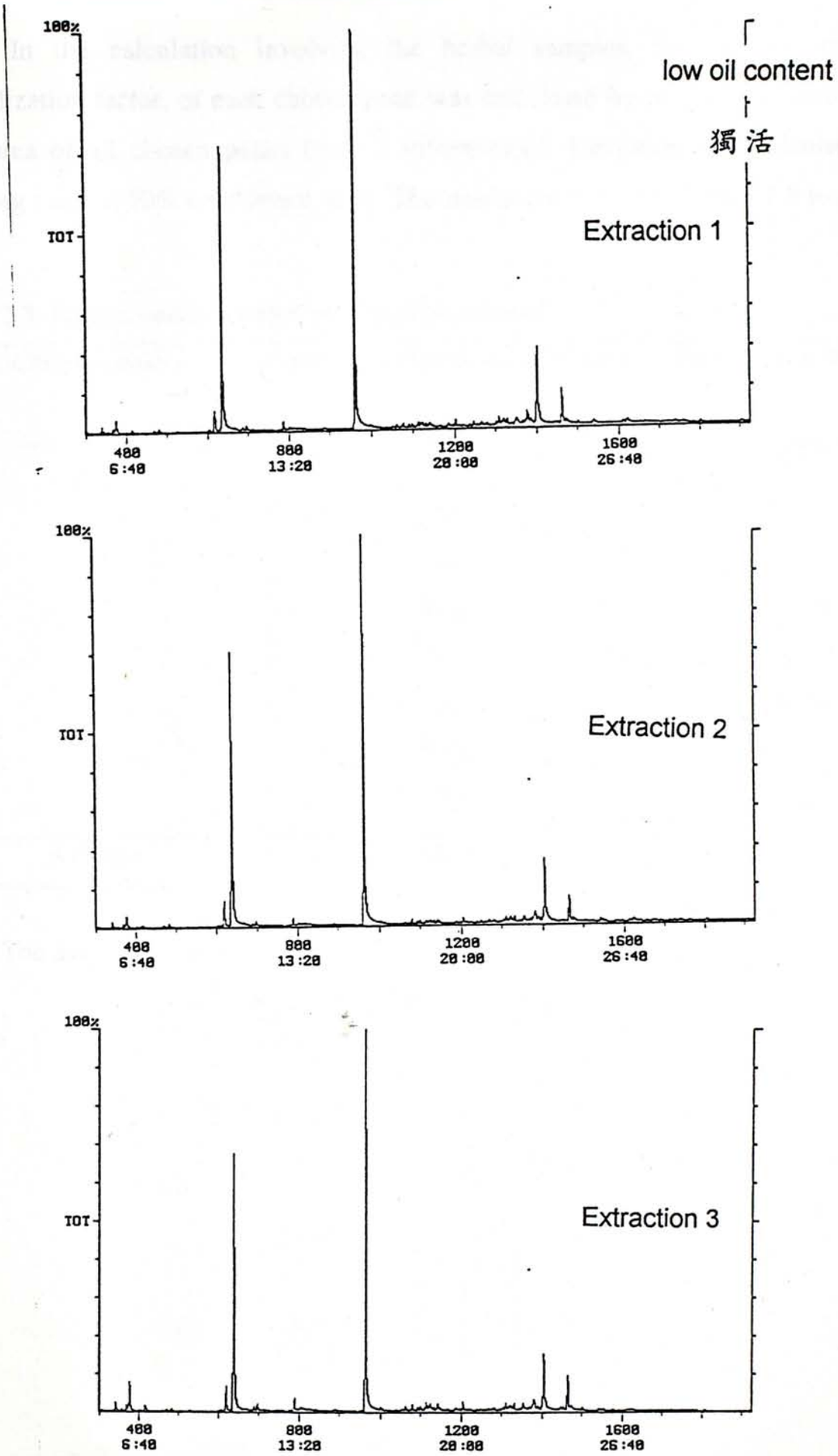


Figure 2.5. Reproducibility of extraction for Duhuo

In the calculation involving the herbal samples, the relative area, or normalization factor, of each chosen peak was calculated by dividing its area to the total area of all chosen peaks (type 2 information). Deviation was calculated by adopting t-test at 90% confidence level. The results are shown in Tables 2.5 to 2.7.

Table 2.5. Jingjie, herbal sample with high oil content

Chosen peaks	Average relative area (%) of 3 extractions	Deviation at 90% CL
1	2.8	0.007
2	0.8	0.003
3	0.8	0.002
4	28.6	0.029
5	2.5	0.008
6	1.1	0.006
7	62.8	0.023
8	0.7	0.002
Average	12.5	0.010

Note: The average relative deviation at 90% confidence level is 8.0%.

Table 2.6. Zisuye, herbal sample with moderate oil contents

Chosen peaks	Average relative area (%) of 3 extractions	Deviation at 90% CL
1	1.2	0.002
2	2.0	0.004
3	34.4	0.072
4	2.5	0.011
5	34.6	0.059
6	2.9	0.005
7	4.7	0.013
8	5.5	0.012
9	13.0	0.116
Average	11.2	0.033

Note: The average relative deviation at 90% confidence level is 29.2%.

Table 2.7. Duhuo, herbal sample with low oil contents

Chosen peaks	Average relative area (%) of 3 extractions	Deviation at 90% CL
1	11.3	0.019
2	11.6	0.013
3	5.4	0.005
4	49.2	0.024
5	19.6	0.023
Average	19.4	0.017

Note: The average relative deviation at 90% confidence level is 8.8%.

From the above calculation, the average relative deviation for the extraction process is about 15%, including the injection errors (which was about 5% when using Finnigan Mat Magnum GC/MS instrument, see 3.4.2.1).

2.4.3 Recovery test

The recovery test was performed by extracting a mixture of reference chemicals. The artificially prepared mixture served as the essential oils of the herbal samples. Without consideration of matrix effect, triplicate extraction of the mixture was carried out using the same experimental procedure and conditions described previously. The extract was diluted to a certain extent with known concentration of internal standards. The percentage recovery was calculated by comparing the area ratios (w.r.t. internal standard ethyl caprylate) of the analytes after extraction with those from the control. The control was prepared by the same experimental procedure, except that the mixture was not added to the reaction flask, but in the collecting organic layer (see Fig. 2.1) under extraction with the analytes in the condensation trap instead.

Results of two recovery tests are reported using the two mixtures described previously, one with lower concentration of reference chemicals to mimic the herbal samples with low essential oil contents (Mixture 1) and another with higher concentration of reference chemicals to mimic the herbal samples with higher essential oil content (Mixture 2). It was noted that the analytes with higher polarities are generally with lower % recovery due to their interactions with water. The results are shown in Table 2.8.

Table 2.8. Recovery of mixture 1 and mixture 2

Reference chemical	% recovery	
	Mixture 1	Mixture 2
α-pinene	105.5	81.6
Myrcene	102.1	91.2
Cineole	101.1	99.8
Fenchone	96.5	97.7
Trans-caryophyllene	70.8	87.0
Humullene	69.3	89.5
Cinnamyl acetate	55.4	90.4

2.5 Discussion

In traditional hydrodistillation or steam distillation methods (section 2.3.1), macroscale analysis is adopted where large amounts of herbal samples (in hundred grams or even kilograms) are used in order to obtain crude volatile oils. In the methodology adopted in this research, by using Dean and Stark-type trap apparatus, the analysis requires only small amounts of plant material (4 g). Besides, further enrichment of the collected essential oils was not required so that analysis by gas chromatography can be carried out after suitable dilution. Thus, the method has the merits of convenience and micro-scale applicability.

Compared with the results on the chemical composition of essential oils in Chinese Medicinal Herbs reported in literature, it was found that the major components extracted in the present study were usually the same as those listed in the literature.

The reproducibility of extraction could affect the choice of the "effective" peaks extracted for building up the library. In different extraction trials, it was found that for the same sample under the same extraction conditions, the chromatographic patterns were similar in the qualitative aspect (retention times). The relative abundance of the components, however, sometimes varied a lot. The calculated relative deviations could be used to estimate the reliability of the extraction process.

From several extraction trials of different reference mixtures and herbal samples, the average relative deviation was about 15% (including injection error). It was found that the relative deviations of the large peaks were about 5%, which were much lower than those for the smaller peaks. Thus, in order to lower the error or increase the reliability of the methodology, peaks with higher abundance should be chosen. Moreover, the deviations calculated from type 1 and type 2 information (see tables 2.3 and 2.4) were close with type 2 being smaller. It was assumed that the relative abundance of individual component had smaller change regardless of the

changes in total content during extraction. Thus, in the analysis of chromatographic patterns discussed in the following chapters, normalization factor (i.e. the ratio of the area of the analyte peak to the total area of chosen peaks, type 2 calculation) was adopted instead of area ratios w.r.t. the internal standard.

The recovery of the mixtures of reference chemicals showed that analytes may not be extracted completely by the adopted extraction apparatus. It may be due to several reasons: (a) some components may still remain in the reaction flask due to their high water affinities or strong interaction with water; (b) some components may be lost on the glassware surface; (c) not all components dissolve in the organic layer; (d) some components may have degraded; and (e) some components may be lost during the collection of the organic layer. The reproducibility should be considerably better for less polar constituents (or components with lower boiling points) [17]. Although the chromatographic patterns may not reveal the actual contents of the components present inside the initial herbal samples, the reproducibility of the chromatograms in the qualitative aspect is satisfactory which was important for recognition. Thus, in the proposed methodology, it could be assumed that the recovery and the degradation of the components were similar, which allowed reasonable comparison to be made.

Chapter 3: Instrumental Analysis of the Essential Oils

3.1 GC analysis

3.1.1 Instrumentation

Gas chromatography is widely used for compound identification in the analysis of mixtures of organic compounds from a variety of chemical classes. The compounds are separated primarily by their volatilities and structures. This allows easy identification and measurement of individual compounds in a sample. In order for a compound to be suitable for GC analysis, it must be evaporated into a vapor state below 400°C and it must not degrade, and must be able to withstand high temperature in the injector port.

A gas chromatograph is essentially a device which enables a small amount of sample to be introduced into an inlet system where it is vaporized and passed into a chromatographic column. To provide suitable conditions for chromatography, the column is held within an oven and a flow of inert carrier gas passes through it. For example, nitrogen was used in our GC/FID instrument. A detector is fitted at the column exit to monitor the separated components as these elute from the column. The detector provides an electrical signal, which is amplified and fed to a recording or data-processing device from which meaningful results can be obtained [24].

A good detection system is one of the most important requirements because the quantitative aspects of analyses depend directly on the detector parameters. Two types of detectors are commonly used in gas chromatographic analysis: flame ionization and mass spectrometric detector. A comparison of various detectors w.r.t. their sensitivities, selectivities, and operational ease indicates that the flame ionization detector (FID) comes closest to being the most sensitive, universal mass detector available for gas chromatography. FID is a common detector because almost all volatile organic chemicals are detected with a response, which does not

significantly depend on the chemical structure of the solutes. The sensitivity to organic solutes varies roughly in proportion to the number of carbon atoms. The FID is an excellent general-purpose detector, particularly with capillary columns, because of its great sensitivity. Its lack of dependence on the flow rate makes it convenient for quantitative measurements. A modern instrumentation for gas chromatograph is shown in Figure 3.1.

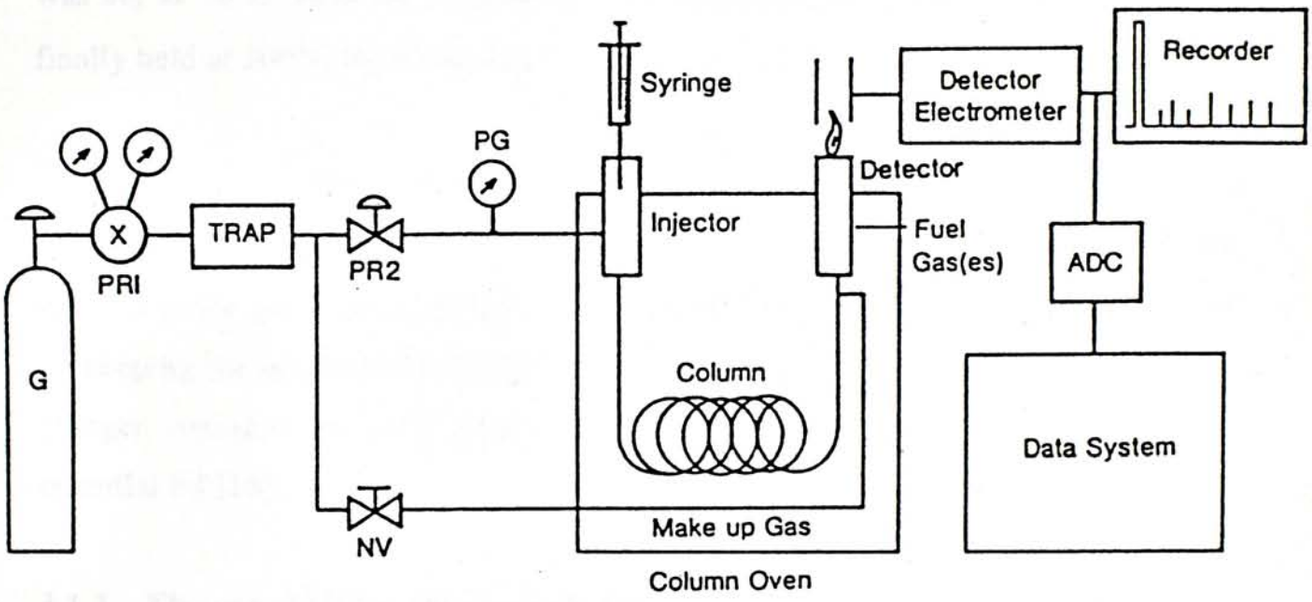


Figure 3.1. Modern instrumentation for Gas Chromatograph

The time that the compound is retained inside the column is called the "retention time". The retention time of each compound is used to qualitatively identify an unknown compound. The peak area is used to measure the amount of the compound. Chromatographic separations only give rise to the retention of solutes, which can never prove what a solute is but only what it is not [25]. The combination of gas chromatography with mass spectrometry can be particularly fruitful for identifying a substance (see section 3.2.3.1)

3.1.2 Instrumental settings

For GC analysis of essential oils extracted from Chinese Medicinal Herbs, the extract was analyzed using an HP-5890 gas chromatograph equipped with a flame ionization detector. A 30m x 0.25mm I.D. HP-5ms (Hewlett Packard) fused silica column was used with nitrogen as carrier gas, with column head pressure of 22 psi. Splitless injector was used with temperature set at 250°C while the FID was set at 280°C. The column temperature was programmed as follows: the initial temperature was set at 70°C, held for 5 minutes, then increased at a rate of 5°C /min, it was finally held at 200°C for 4 minutes. The total analysis time was 35 minutes.

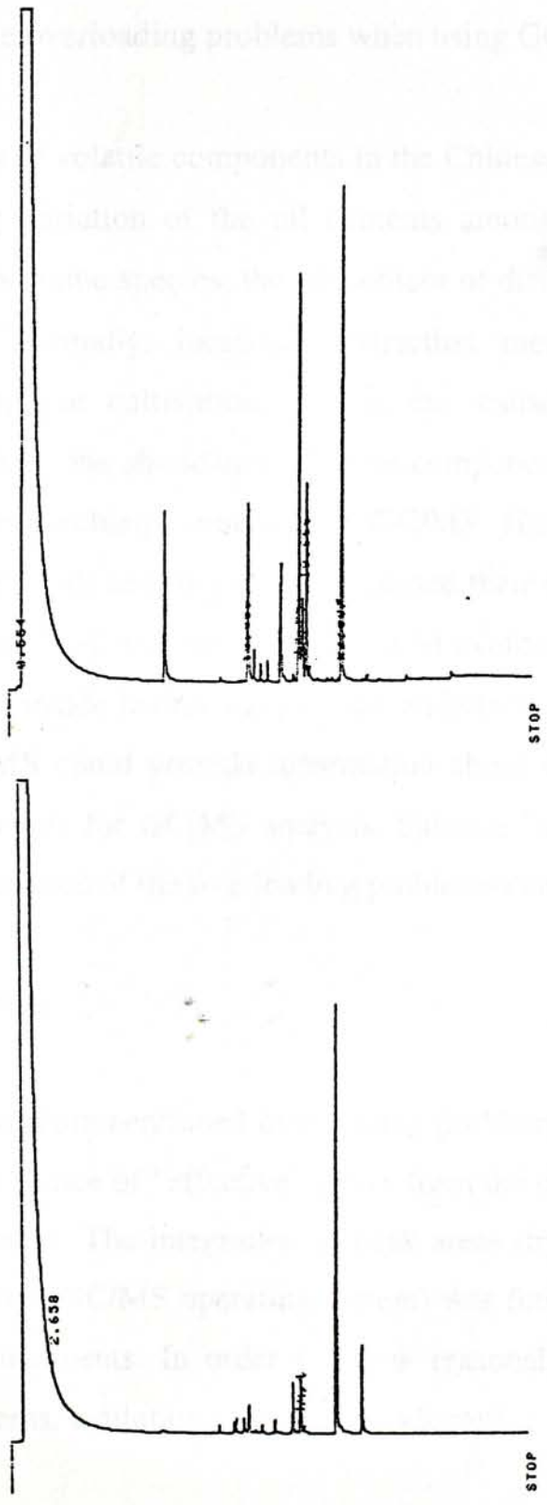
To analyze the essential oil, 0.5 µl the diluted extract (the crude extract was diluted 10 fold with *n*-hexane with 45 ppm internal standards, see section 2.3.2) was injected in the gas chromatograph. The hot needle injection technique was employed by keeping the needle in the heated injector device for 3 seconds before pushing the plunger, owing to the wide range of volatility of the components in the extracted essential oil [16].

3.1.3 The use of GC in the analysis of essential oils

3.1.3.1 Qualitative data

GC was used to preliminarily estimate the number of components in the extracted oil. Capillary column gas chromatography can generate retention indices very precisely, assuming standardization of stationary film polarity, carrier gas flow-rate, stationary phase film thickness and temperature programming rate [18]. When comparing the chromatograms of the extract and the reference chemicals under the same instrumental conditions, the identities of the components could be guessed by matching the retention times, provided that the corresponding chromatograms of reference chemicals were available. However, the correct identities of the components could not be confirmed by only comparing the retention indices (MS

provide complementary data for the confirmation of the identities of the components). Nevertheless, the retention times of the chromatograms could be compared to estimate their similarity. Some examples are illustrated in Figure 3.2.



白朮

蒼朮

Figure 3.2 Comparison between two herbal samples, Baizhu and Canzhu, from GC patterns

3.1.3.2 Quantitative data

In quantitative aspect, gas chromatography was used to preliminarily estimate the relative abundance of the components in the extracted oil from the peak areas. It can help to prevent the overloading problems when using GC/MS for analysis.

In the analysis of volatile components in the Chinese Medicinal Herbs, it was discovered that great variation of the oil contents among different species could occur. Even among the same species, the oil content of different samples could vary in a certain extent. Actually, locations, extraction methods, sources, time of cultivation and method of cultivation, season, etc. cause changes in the aroma composition. Sometimes, the abundance of some components could be so large that may cause overloading problems when using GC/MS. However, overloaded peaks affected the resolution of all nearby peaks and caused their retention times to deviate from the normal values [14], and hence it should be avoided. Owing to the fact that the actual oil contents inside herbal samples are uncertain, analysis of essential oils by GC prior to GC/MS could provide information about the oil contents, suitable dilution can then be made for GC/MS analysis. Suitable "analysis window" can be obtained and the occurrence of the overloading problems can be prevented.

3.1.3.3 Dilution strategy

Apart from the aforementioned overloading problem, the variation of the oil content can affect the choice of "effective" peaks from the chromatographic patterns for database development. The integration of peak areas (by the built-in integration program provided by the GC/MS operating system) was found to be affected by the abundance of the components. In order to allow reasonable comparison between chromatographic patterns, a dilution strategy was adopted.

The extent of dilution of oil extracts depended on the area ratio of the most abundant peak to the internal standard (using ethyl caprylate). After dilution, the

concentrations of the components (including internal standards) were controlled below 100 ppm with the most abundant peak to be at least 40 ppm. The extent of dilution depended on the value, R, defined in Equation 3.1. The dilution factors are listed in Table 3.1.

R = (area of the most abundant peak) / (area of internal standard) (Eqn. 3.1)

Table 3.1. Dilution factors

Range of R	Dilution factor
$2 < R \leq 5$	2.5
$5 < R \leq 10$	5
$10 < R \leq 20$	10
$20 < R \leq 40$	20
$40 < R \leq 60$	25
$60 < R$	50

However, for some samples, the oil contents may be so low that the area of the largest peak was less than 10% of that of the internal standard. For such cases, the concentration of the oil has to be increased 5 times before injection instead.

Some chromatograms are shown in Figure 3.3 to illustrate the chromatograms of samples after dilution. It can be noted that the patterns are more comparable.

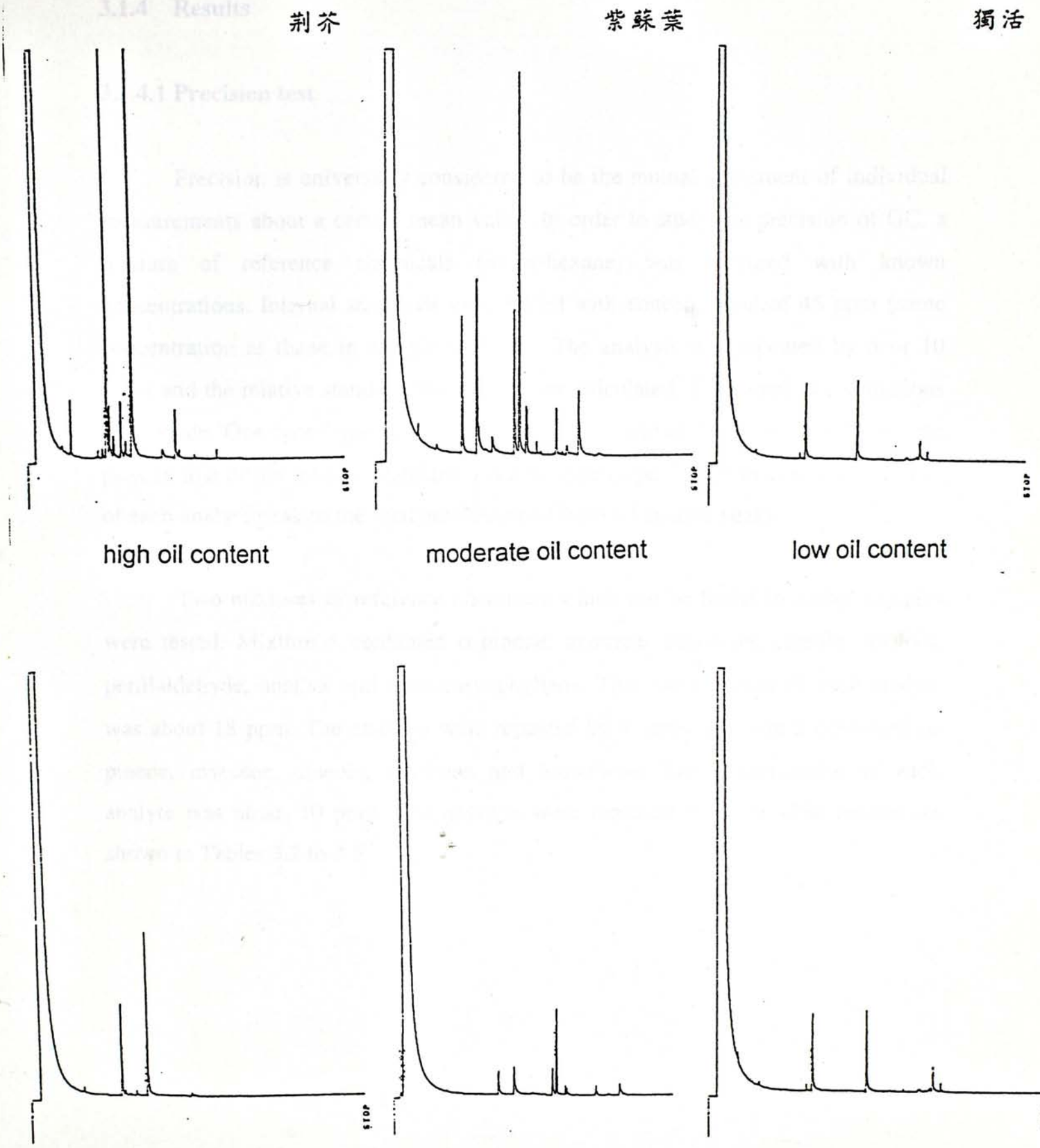


Figure 3.3. Chromatograms of samples after dilution

3.1.4 Results

3.1.4.1 Precision test

Precision is universally considered to be the mutual agreement of individual measurements about a certain mean value. In order to study the precision of GC, a mixture of reference chemicals (in *n*-hexane) was prepared with known concentrations. Internal standards were added with concentration of 45 ppm (same concentration as those in sample analysis). The analysis was repeated by 5 or 10 times and the relative standard deviations were calculated. Two types of calculations were made. One type (type 1) was to calculate the ratio of the areas of each analyte peak to that of the internal standard. Another type (type 2) was to calculate the ratio of each analyte peak to the total area obtained from all analyte peaks.

Two mixtures of reference chemicals which can be found in herbal samples were tested. Mixture 1 contained α-pinene, myrcene, limonene, cineole, linalool, perillaldehyde, anethol and trans-caryophyllene. The concentration of each analyte was about 18 ppm. The analysis were repeated by 5 times. Mixture 2 contained α-pinene, myrcene, cineole, fenchone and humullene. The concentration of each analyte was about 10 ppm. The analysis were repeated 10 times. The results are shown in Tables 3.2 to 3.5.

Table 3.2. Mixture 1, type 1 information

Reference chemicals	Mean peak area ratio (%)	Relative standard deviation* (%)
α-pinene	87.2	1.6
Myrcene	54.1	1.7
Limonene	87.5	2.0
Cineole	106.0	2.7
Linalool	56.8	0.8
Perillaldehyde	51.4	1.2
Anethol	79.8	1.6
Trans-caryophyllene	101.4	2.0

* The average relative standard deviation was 1.7%.

Table 3.3. Mixture 1, type 2 information

Reference chemicals	Mean peak area ratio (%)	Relative standard deviation* (%)
α-pinene	14.0	0.48
Myrcene	8.7	0.47
Limonene	14.0	0.54
Cineole	17.0	1.10
Linalool	9.1	1.46
Perillaldehyde	8.2	1.47
Anethol	12.8	0.33
Trans-caryophyllene	16.2	1.17

* The average relative standard deviation is 0.88%.

Table 3.4. Mixture 2, type 1 information

Reference chemicals	Mean peak area ratio (%)	Relative standard deviation* (%)
α-pinene	23.0	4.0
Myrcene	14.2	3.0
Cineole	21.3	2.6
Fenchone	20.4	2.6
humullene	24.0	2.4

* The average relative standard deviation was 2.9%.

Table 3.5. Mixture 2, type 2 information

Reference chemicals	Mean peak area ratio (%)	Relative standard deviation* (%)
α-pinene	17.6	4.3
Myrcene	10.9	5.0
Cineole	16.3	2.3
Fenchone	15.6	2.2
humullene	18.4	2.4

* The average relative standard deviation was 3.2%.

It was found that the precision of the instrument as indicated by the relative standard deviations of 5 or 10 replicate analysis is good. The relative standard deviation calculated for mixture 2 was higher than that from mixture 1, possibly because of the lower concentrations of reference chemicals in mixture 2, thus enlarging the relative deviations.

3.1.4.2 Linearity

The ability of an instrumental analysis technique to handle a wide range of analyte concentrations means that the results are calculated, and the random errors evaluated, in a particular fashion that the recorded signals are in a linear relationship with the change of concentrations. The usual procedure is as follows. The analyst takes a series of samples with known concentrations of the analytes. These calibration standards are measured in the analytical instrument under the same conditions as those used for the test samples [26, p.102].

The response factors of different components acted on the column and the detector are different. Some components with small response factors have slight increase in peak areas upon increase in concentrations, or vice versa. Thus, the ratio, *R* (defined in equation 1), could not reveal the actual concentration of the components in the extract. For those with low *R* values, it may be due to the actual low abundance of these components, or it may be due to the low response factors despite the high concentration, or the opposite was true. Thus, the largest peak may have a higher concentration than expected. Studying linearity of GC and GC/MS could help to prevent the possibility in column saturation. This helped to confirm the validity of the dilution strategy. The ratio of the components should not be affected upon dilution theoretically, thus the choice of "effective" peaks (see section 4.2.1) for database development was not affected.

In order to study the linearity of the concentration range, a mixture of reference chemicals, together with internal standards, with known concentrations were prepared. The mixtures were analyzed by GC under the same conditions as those used for the testing samples. A series of mixtures with different concentrations were examined.

The tested analytes were α-pinene, cineole fenchone, trans-caryophyllene, cinnamyl acetate and hexadecane, which can be found in herbal samples. The

concentrations of the analytes were 5, 10, 20, 40, 60, 80, 100, 150, 200, 500 ppm with internal standards of about 45 ppm. The area ratios of each analyte to that of the internal standard were calculated. The calibration graphs with area ratios against concentrations were plotted and shown in Figure 3.4. The individual calibration graph is shown in Appendix A. The results are tabulated in Table 3.6 below.

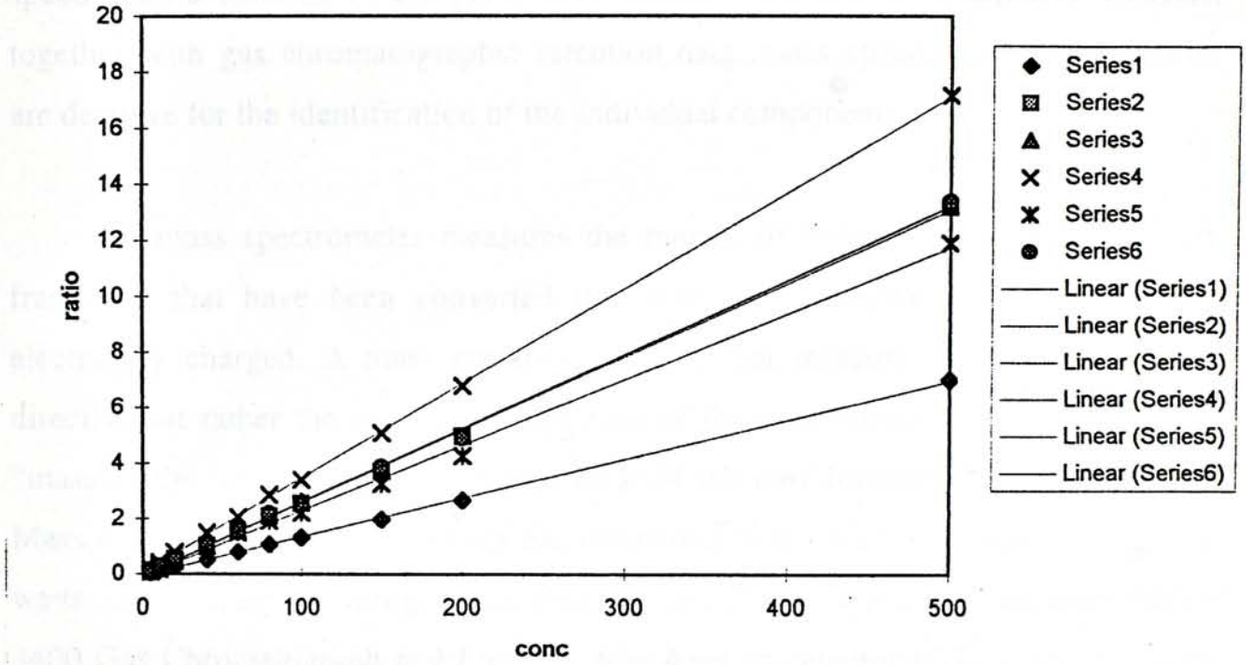


Figure 3.4. Calibration graphs for the reference chemicals obtained by GC

Table 3.6. Linearity of reference chemicals

Analyte	Series	Concentration range (ppm)	Correlation coefficients
α-pinene	1	5-500	0.9995
Cineole	2	5-500	0.9992
Fenchone	3	5-500	0.9991
Trans-caryophyllene	4	5-500	0.9998
Cinnamyl acetate	5	5-500	0.9975
hexadecane	6	5-500	0.9988

3.2 GC/MS analysis

3.2.1 Instrumentation

The instrumentation of gas chromatography in section 3.1.1 can also be applied here except that a mass spectrometer is used as the detector. The mass spectrometric detector is a sophisticated detection system. In qualitative analysis, together with gas chromatographic retention data, mass spectrometric information are decisive for the identification of the individual components.

A mass spectrometer measures the masses of individual molecules and its fragments that have been converted into ions, i.e. molecules which have been electrically charged. A mass spectrometer does not measure the molecular mass directly, but rather the mass to charge ratio of the ions formed. It is common that "mass of the ion" is referred because most of the ions formed are singly-charged. Mass spectrometry combined with gas chromatography is clearly superior in many ways. In this study, a Finnigan Mat Magnum GC/MS system was used (with Varian 3400 Gas Chromatograph and Finnigan Mat ion trap detector (ITD)), and its block diagram is shown in Figure 3.5.

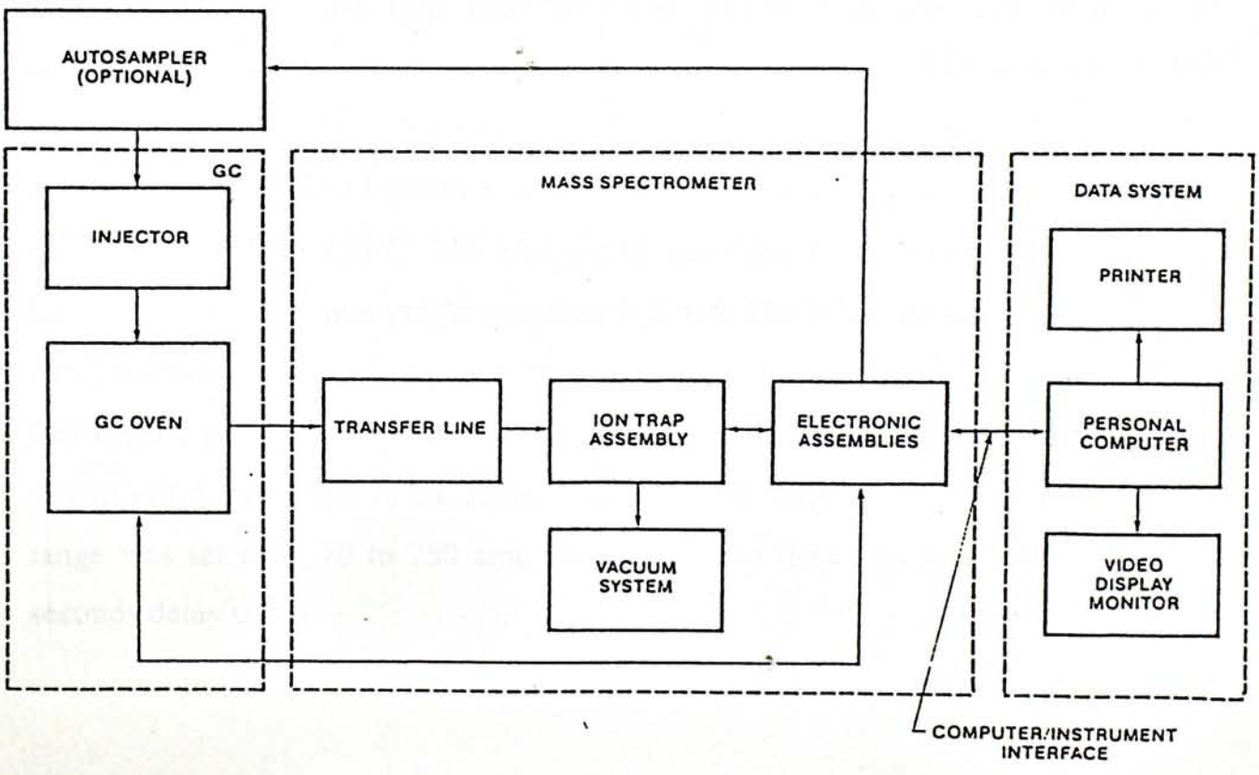


Figure 3.5. The block diagram of Magnum

In the ion trap detector, ions of similar charge are sorted inside the mass analyzer based on their mass-to-charge ratios. The lighter ions come out of the analyzer first. It operates on a principle similar to the quadruple. However, rather than allowing ions to pass through it, the ion trap can store ions for subsequent experiments. The ions are first trapped in an enclosed ion trap chamber, and then ions are sorted out by changing the electric field inside the trap. The ions are then sequentially ejected out of the trap by increasing mass-to-charge ratios for ion detection [23].

The mass spectra generated by fragmenting a molecule can be used to elucidate a chemical structure of an unknown compound. With the availability of the compiled mass spectra of known compounds, it is possible to determine the identity of the unknown compound or to determine the chemical structure of the unknown.

3.2.2 Instrumental settings

After analysis of the oil extract by GC and with suitable dilution, the extract was analyzed by GC/MS. The instrument used was Finnigan Mat Magnum GC/MS Systems (Varian 3400 Gas Chromatograph, Finnigan Mat Magnum Mass Spectrometer with ion trap detector). The extract was analyzed with a low background DB-5ms fused silica capillary column, 30m x 0.25 mm I.D. (J&W Scientific, USA). Helium was used as the carrier gas with column head pressure set at 12 psi. Split/Splitless injector was used with temperature set at 250°C. The transfer line was also set at 250°C. The MS source manifold temperature was 220°C. The column temperature was programmed as follows: The initial temperature was set at 70°C and was held for 2 minutes, then increased at a rate of 5°C /min and held at 200°C for 4 minutes. The total analysis time was 32 minutes. The MS parameters are shown as follows: electron ionization was used with ionizing voltage 70 eV; the mass range was set from 70 to 250 amu; the acquisition time was 32 minutes with 300 seconds delay time.

To analyze the essential oil, 0.5 µl of the diluted extract was injected with the hot needle injection technique by keeping the needle for 3 seconds in the heated injector device before pushing the plunger.

3.2.3 The use of GC/MS in the analysis of essential oils

3.2.3.1 Identification by GC/MS

The identification by GC/MS may be possible with 10 ng or more of a compound. The identification or matching of the chromatographic peaks is based on the comparison of the mass spectra and GC retention times. This latter property can be of considerable value, especially when the capillary columns of very high resolving power are used. Although the resolution of such columns is not sufficient to give an identification solely based on the retention time, analytes having different retention times are clearly not identical compounds. The value of mass spectra lies in their high information content. Most compounds can be distinguished on the basis of their mass spectra, and good matches give a high confidence of identity. Thus, MS data can serve as an auxiliary tool to distinguish chromatographic peaks for comparison, even for the peaks which are close together. Overall, the chromatograms obtained by GC/MS could be used for database development after analysis of the patterns. The analysis was based on the comparison of the "effective" peaks (see section 4.2.1). The retention indices of the chromatographic peaks served as codes for the herbal samples in qualitative aspect.

3.2.3.2 Abundance information

In quantitative aspect, by making use of the total ion chromatograms obtained by GC/MS, the area ratios or normalization factors of the "effective" peaks could be calculated. The total mass chromatogram is the total signal measured by the mass spectrometer during a GC/MS run. This chromatogram is similar to a typical gas chromatogram [24].

Since GC/MS has a lower detection limit than GC/FID, components with smaller abundance could be spotted out.

3.2.4 Results

3.2.4.1 Precision

In order to study the precision of GC/MS, the method described in section 3.1.4.1 was applied except that a GC/MS was tested instead of a GC. A mixture of reference chemicals (in *n*-hexane) was prepared with known concentrations. Internal standards were added with concentration of 45 ppm (same concentration as those in sample analysis). The data were obtained from 5 or 10 replicate runs. The ratios of each analyte peak to the total area obtained for all analyte peaks were calculated.

Two mixtures of reference chemicals were tested. Mixture 1 contained myrcene, limonene, cineole, linalool, perillaldehyde, anethol and trans-caryophyllene. The concentration of each analyte was around 18 ppm. The data were obtained from 5 replicate runs. Mixture 2 contained myrcene, cineole, fenchone, trans-caryophyllene and humullene. The concentration of each analyte was around 10 ppm. The data were obtained from 10 replicate runs. It was found that the precision, as indicated by the relative standard deviations of the measurements, of the instrument was good. The results are shown in Tables 3.7 to 3.10.

Table 3.7. Precision data for mixture 1, type 1 information

Reference chemicals	Mean peak area ratio (%)	Relative standard deviation* (%)
Myrcene	122.8	2.4
Limonene	74.6	4.4
Cineole	88.0	1.0
Linalool	53.9	2.8
Perillaldehyde	41.4	4.3
Anethol	79.4	5.2
Trans-caryophyllene	96.6	2.2

* The average relative standard deviation was 3.2 %.

Table 3.8. Precision data for mixture 1, type 2 information

Reference chemicals	Mean peak area ratio (%)	Relative standard deviation* (%)
Myrcene	22.1	1.7
Limonene	13.3	2.1
Cineole	15.8	5.2
Linalool	9.6	2.1
Perillaldehyde	7.4	1.7
Anethol	14.3	3.1
Trans-caryophyllene	17.4	4.3

* The average relative standard deviation is 2.8 %.

Table 3.9. Precision data for mixture 2, type 1 information

Reference chemicals	Mean peak area ratio (%)	Relative standard deviation* (%)
Myrcene	9.2	2.7
Cineole	22.0	3.2
Fenchone	21.1	2.6
Trans-caryophyllene	24.6	11.8
Humullene	27.7	4.9

*The average relative standard deviation was 5.0%.

Table 3.10. Precision data for mixture 2, type 2 information

Reference chemicals	Mean peak area ratio (%)	Relative standard deviation* (%)
Myrcene	8.8	5.1
Cineole	21.0	3.6
Fenchone	19.7	2.9
Trans-caryophyllene	23.7	7.9
Humullene	26.8	2.8

* The average relative standard deviation was 4.5%.

3.2.4.2 Linearity

It had to be verified that the concentration range after dilution (highest concentration was about 100 ppm) was linear when using GC/MS to avoid column saturation.

In order to study the linearity of the concentration range, a mixture of reference chemicals with known concentrations (together with the internal standards) were prepared. The mixtures were analyzed with the GC/MS under the same

conditions as those used for the samples. A series of mixtures with different concentrations were used.

The tested analytes were cineole, fenchone, trans-caryophyllene and hexadecane, which can be found in herbal samples. The concentrations of the analytes were 0.5, 1, 5, 10, 20, 40, 60, 80, 100 ppm with internal standards of around 45 ppm. These calibration standards were analyzed with the GC/MS under the same instrumental conditions described above. The area ratios of the area of each analyte to that of the internal standard were calculated. It was found that the concentration range under study was linear. The calibration graphs with area ratios against concentrations were plotted and shown in Figure 3.6. The individual calibration graph was shown in Appendix B. The data for the calibration graphs and correlation coefficients are shown in Table 3.11 below.

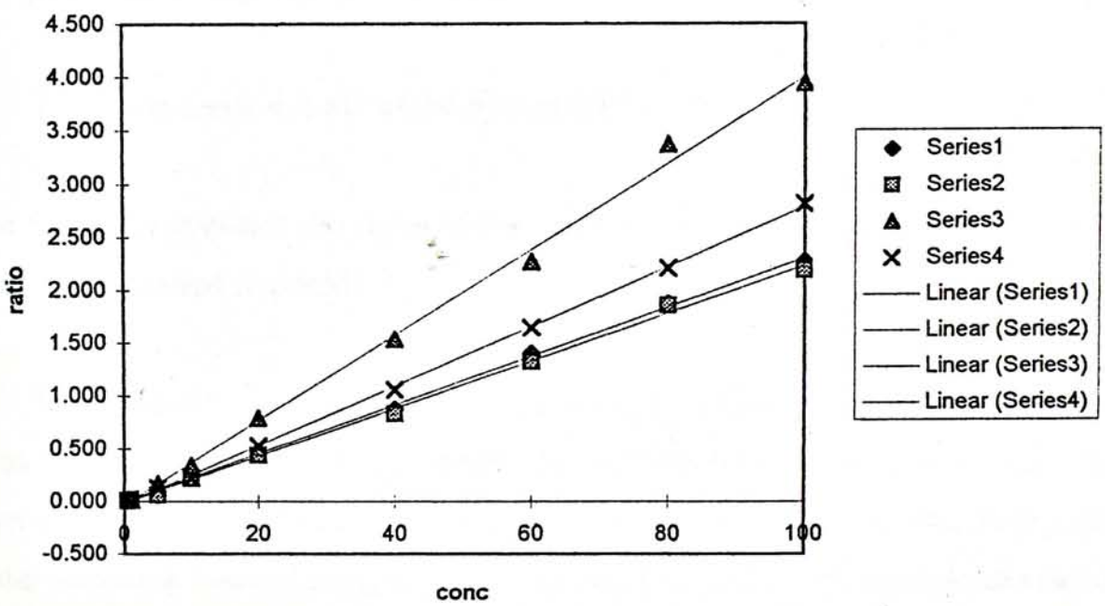


Figure 3.6. Calibration curves of the reference chemicals obtained by GC/MS

Table 3.11. The linear calibration range for the reference chemicals

Analyte	Series	Concentration range (ppm)	Correlation coefficients
Cineole	1	0.5-100	0.9986
Fenchone	2	0.5-100	0.9979
Trans-caryophyllene	3	0.5-100	0.9972
hexadecane	4	0.5-100	0.9994

3.2.4.3 Detection limit

In general terms, the limit of detection of an analyte may be described as that concentration which gives an instrument signal significantly different from the 'blank' or 'background' signal [26, p.115].

Detection limit can be obtained by calculating the standard deviation of several trials from the background at the analyte position. It can be calculated using Equation 3.2 according to the IUPAC definition [29].

Detection Limit = 3 SD of the background at the analyte position (Eqn. 3.2)

where SD is the standard deviation of the ratio of the area of the background signal to that of the internal standard.

Since the components of the essential oil of different herbal samples could not be exclusively analyzed to obtain the corresponding detection limits, four analytes were studied to estimate their detection limits. A solution containing two internal standards (around 45 ppm) were injected five times. The standard deviations of ratio of the area of the background signal of each analyte to that of the internal standard were calculated. From the calculated results, it can be seen the detection limits of the analytes are in the range of 0.01 to 0.03 ppm. The results are shown in Table 3.12.

Table 3.12. Detection limits of some analytes present in essential oils

Analyte	Mean area ratio of background to IS	Standard deviation of area ratio (SD)	3SD	Detection limit (ppm)
Cineole	0.00015	0.00011	0.00032	0.01
Fenchone	0.000087	0.00014	0.00041	0.02
Trans-caryophyllene	0.00016	0.00015	0.00045	0.01
hexadecane	0.00071	0.00027	0.00082	0.03

3.2.4.4 Chromatographic patterns of herbal samples obtained by GC/MS

The chromatograms for data analysis and database development were obtained by GC/MS. The corresponding chromatograms of 79 samples (treated as known samples) used for library development are shown in Appendix C.

3.3 Discussion

In GC analysis, such useful information as the number and abundance of components in the extracted oil can be obtained. By referring to the retention indices, the chromatographic peaks can be compared qualitatively. However, the correct identities of the components cannot be confirmed just by comparing the retention indices. With the combination of an MS detector, the mass corresponding to each peak in different chromatographic patterns can be obtained and compared.

In quantitative aspect, peak areas (which were used in this study) or peak heights, can provide abundance information of individual components. It was found that the major components extracted from the herbal samples were usually those reported in the literature. Together with qualitative data, obtained from the chromatographic patterns of different samples, preliminary ideas on the similarity or difference of the herbal samples could be drawn. Besides, the abundance information could provide an estimation of the oil content in the herbal samples. Overloading

problems could therefore be avoided when using GC/MS where ion saturation could occur.

Owing to the variation in oil contents, dilution strategy was adopted to obtain suitable "analysis window" for different herbal samples to prevent overloading of the GC or GC/MS. It was noted that the chromatographic patterns were more comparable among the same species or between different species. Moreover, different species could be analyzed by using the same standardized procedure. These are important considerations for database development. From the linearity study, it was shown that the concentration range under study was acceptable. This helped to ensure the validity of the procedure in analyzing the chromatographic patterns.

From the results of detection limit, it was found that the detection limits of the analytes could be down to around 0.02 ppm. In the analysis procedure for choosing the "effective" peaks for database development, by setting the threshold value at 1.50 or even 0.50 (see section 4.2.1), the peaks from analytes at around 0.1 ppm were not included in the integration procedure. Thus, the analysis was unlikely affected by the background interference.

In this study, normalization factor, i.e. ratio of area of the analyte peak to the total area of all chosen analyte peaks, was used instead of the absolute abundance or the area ratios of analytes to internal standards. The normalization factor was adopted due to the variation of total oil contents in herbal samples from different sources, even of the same species. This may be due to the differences in storage and cultivated methods, etc. However, the relative amounts of the components in herbal samples, especially the major constituents, usually vary to a smaller extent.

Furthermore, it was found that the flame ionization detector was more stable than the mass spectrometric detector. It could be shown by the higher precision of GC analysis. It may bedue to the randomness of the ions formed in the ion source.

In summary, through instrumental analysis, comparable chromatographic patterns could be obtained by GC analysis, dilution strategy and GC/MS analysis. Although the chromatographic patterns may not reveal the actual oil content inside the herbal samples, it can be assumed that the recovery and the degradation of the components are similar which makes recognition possible. These chromatographic patterns have some characteristic features for a particular herbal species. Furthermore, chromatographic peaks can be compared or matched based on the retention indices and mass spectra information, which can help to assess the possibility in building up a scheme for recognition.

Chapter 4: Development of a system for recognition

4.1 Introduction

For the recognition of the Chinese Medicinal Herbs under study, a database containing the information from the chromatographic patterns (obtained by the aforementioned instrumental analysis) was developed. Excel, which is a user-friendly spreadsheet software, has been used for the proposed method. Excel contains many built-in functions which facilitate the calculations, and customized programs can be written using visual basic languages, making future modification of the methodology possible.

The proposed system comprised the library and matching sections. Prior to the data entry, the chromatographic patterns from GC/MS need to be analyzed. "Effective" peaks (for definition, see 4.2.1) were extracted from the patterns using some proposed criteria, which were based on the abundance of the components. The qualitative and quantitative indices assigned to the chosen peaks were then calculated. They were referred to as the relative retention indices (w.r.t. the two internal standards) and the normalization factors (for definition, see 4.3.2) respectively. From the instrumental analysis of the essential oils, it was found that though the composition of the oil varied among the same species of the herbal samples, the major peaks of their chromatographic patterns are usually similar to a certain extent. Therefore, they were extracted and used to constitute the "characteristic" peaks (for definition, see 4.2.2) for a particular kind of herbal drug.

The library section of the system contains information of the "effective" and "characteristic" peaks of the herbal samples. Their retention indices and normalization factors are calculated here. Besides, the peaks can be sorted or arranged for the matching section, where comparison with the information of the unknown takes place. The information of the documented samples can be retrieved easily.

In the matching section, an "effective" peak and "characteristic" peak matching method is developed through the proposed calculation algorithms. On the basis of the matched result, the similarity score between the unknown sample and target herbal drug is calculated. The matching section consists of two parts: one involves the comparison with the "characteristic" peaks of the herbal drugs, mainly in the qualitative aspect (i.e. relative retention times). Another involves the comparison with the whole picture of "effective" peaks of each individual species, including both the qualitative (relative retention times) and quantitative data (normalization factors).

Apart from the library and the matching sections, the database contains the input and output layouts. The input layout is used to enter the unknown data for calculation. The output layout shows the matching results and some information, such as the retail names, about the searched target drug.

Some assumptions and rules were proposed to indicate the scope or the limit of the applicability of the proposed database, which will be discussed below.

4.2 Analysis of chromatographic patterns

The manner of the chromatographic analysis has a profound effect on the efficiency of recognition. The two most important criteria are the intensity and specificity of the selected index peaks. These peaks should be among the most intense components so that they can be recognized easily in a complex mixture. They should also be relatively specific to a particular oil such that the chance for another oil containing the same index peaks is smaller [11]. The extraction of "effective" peaks in this study satisfies the former criterion while the selection of "characteristic" peaks satisfies both criteria.

4.2.1 Extraction of "effective" peaks

The definition of "effective" peaks used in this study are the peaks, which fulfill the proposed criteria (mainly based on abundance consideration), extracted from the chromatographic patterns and used for further comparison and calculation. They should not be affected by interference from by the background and impurities. They should come from the samples, and usually those components with higher abundance. The extraction of "effective" peaks from the chromatograms depends on the relative abundance and the separation of the peaks.

Upon suitable dilution, the concentrations of the components used for GC/MS analysis should be under 100 ppm as previously described. However, there should be some components (including the internal standards) with concentrations at least about 40 ppm. Furthermore, the range of retention times is from 300 to 1920 seconds. These constitute the "analysis window" shown in Figure 4.1.

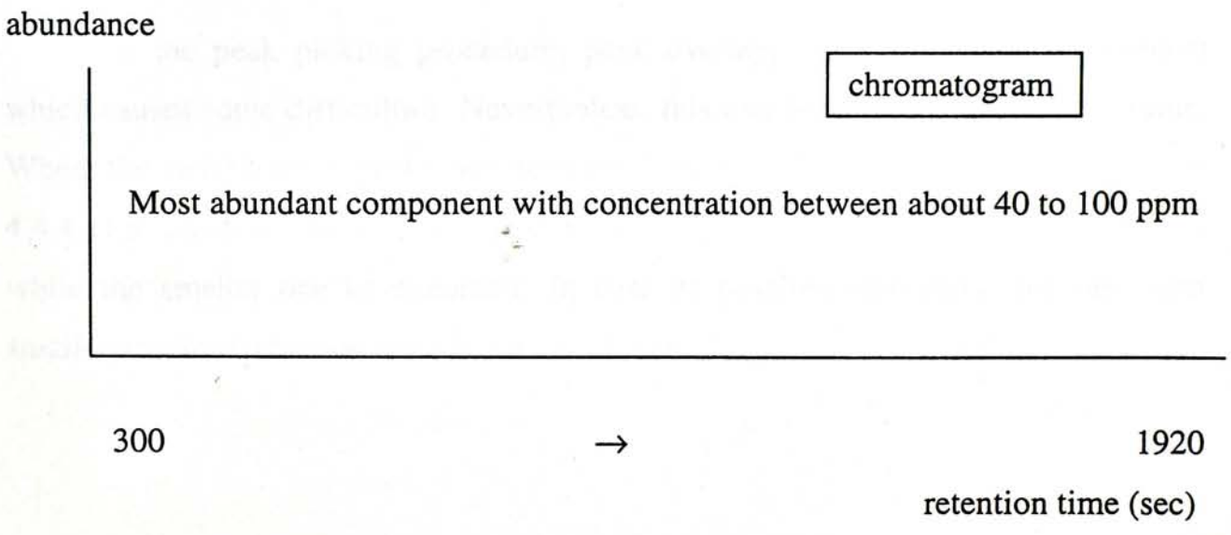


Figure 4.1. The "analysis window" of GC/MS chromatograms

In the process of extracting "effective" peaks, the major peaks must be present in the "analysis window" for the auto-integration process to take place. This auto-integration process is a built-in function in the GC/MS system. The integration is based on the threshold value so that the integration results are affected by the abundance of the largest peak in the "analysis window". With this consideration, the procedure for "effective" peak extraction is described below:

(a) The major peaks must be present in the "analysis window" for peak integration.
(b) Threshold value is set at 1.50.
(c) If there are less than 10 most abundant peaks (excluding those from internal standards) being integrated, the threshold value is set at 1.00 to see whether at least 10 most abundant peaks can be chosen.
(d) If not, the threshold value is set at 0.50 and the integration is performed again. Threshold value will not be reduced further even if the number of chosen peaks are still less than 10.
(e) After spotting the peaks, integration is performed at threshold value of 1.50 to calculate the peak areas. Manual integration may be needed for some peaks.

In the peak picking procedure, peak overlapping is an inevitable problem which causes some difficulties. Nevertheless, this can be resolved by making a rule. When the neighboring peaks are separated by less than 6 seconds (see section 4.4.4.1) in the absolute retention time scale, the peak with the larger area is chosen while the smaller one is discarded. In case of possible ambiguity, the one with smaller absolute retention time is always chosen. The rule is illustrated in Figure 4.2.

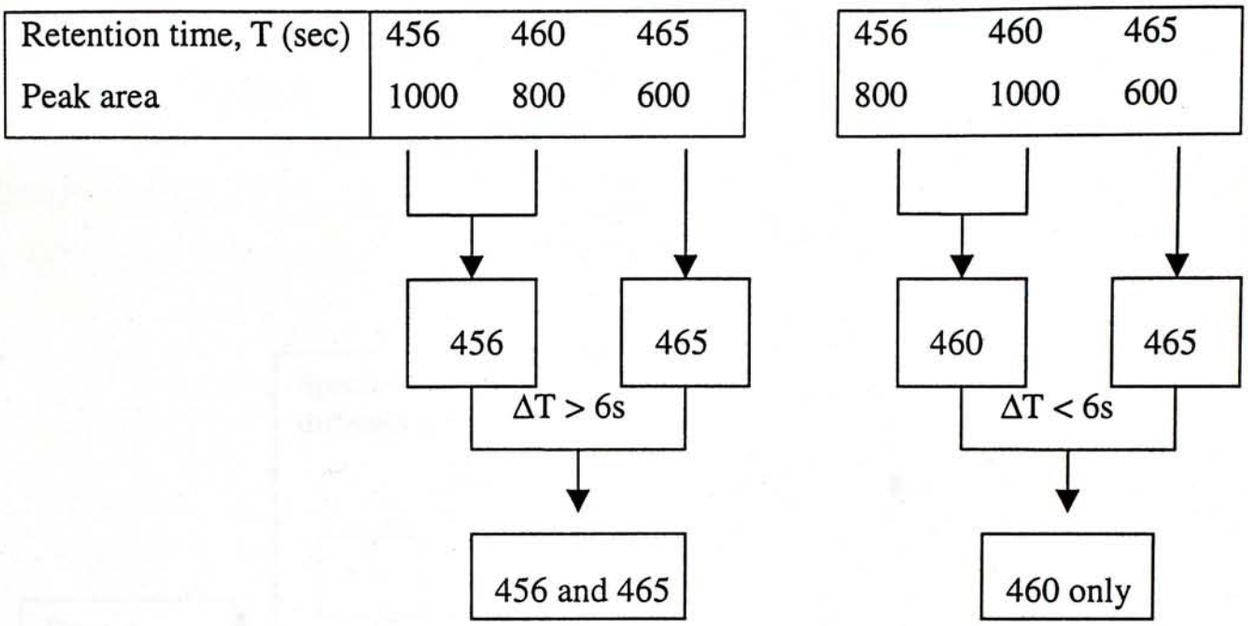


Figure 4.2. The rule of picking peaks when peak overlapping is encountered.

4.2.2 Extraction of "characteristic" peaks

The aforementioned "effective" peaks are those chosen from the chromatographic patterns of individual herbal species which satisfy the proposed criteria of abundance. The "characteristic" peaks discussed here are those which constitute the specific features for a particular herbal drug. Here, "characteristic" peaks are chosen by extracting the common peaks from the "effective" peaks which exist in the same herbal drugs but from different sources. Ten most abundant "effective" peaks are used for the extraction.

In the process of extracting the "characteristic" peaks, when the herbal drug from more than half of the sources (over half of the occurrence probability) contain a particular peak, this peak is considered to be one of the "characteristic" peaks. A diagram shown in Figure 4.3 is used for clarification. Two examples are illustrated in Table 4.1 and 4.2. The data shown in the two tables are relative retention indices discussed in section 4.3.1 below.

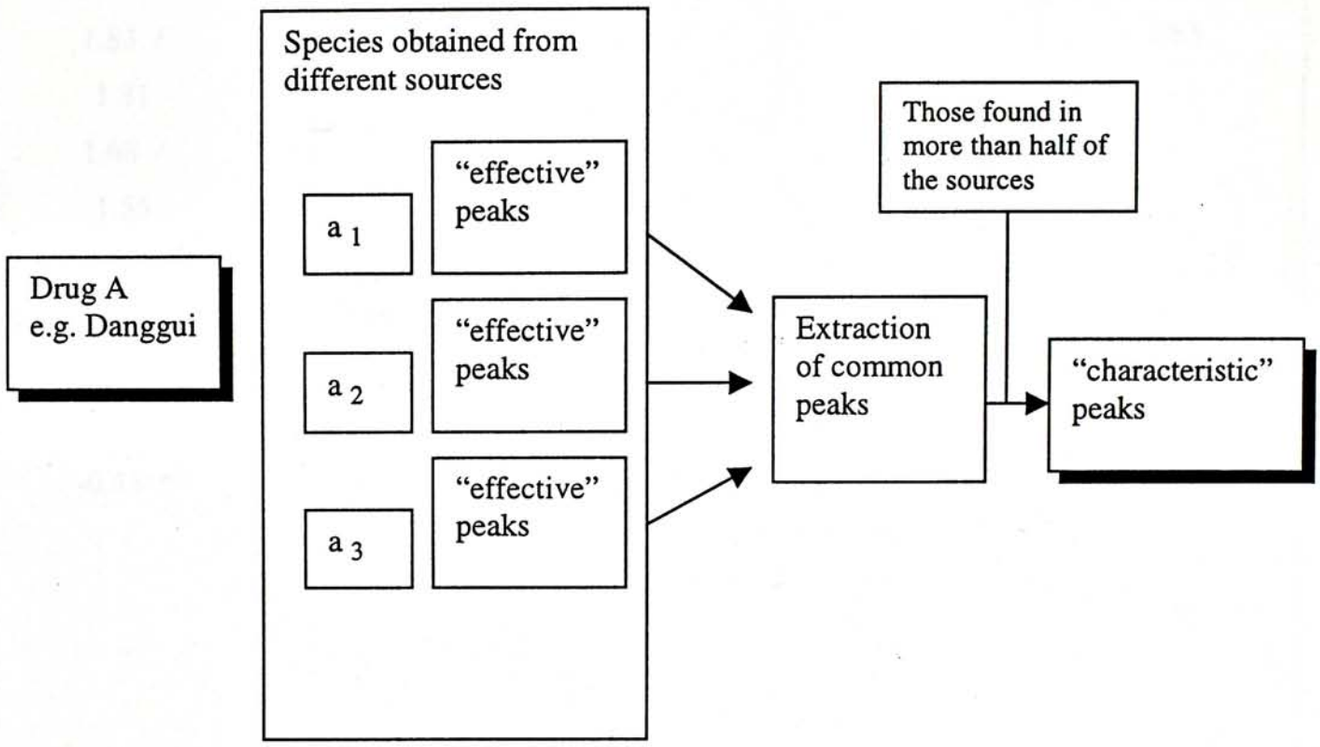


Figure 4.3. Extraction of "characteristic" peaks from the "effective" peaks of species from different sources.

Table 4.1. Extraction of "characteristic" peaks from the "effective" peaks. Sample: Qianghuo

10 (or less) effective peaks in relative retention times, *t* *				Characteristic peaks
Sources of samples #				
Hong Kong	Beijing	Guangzhou	Reference	
1.90	-	1.90	-	-
-	1.86	-	-	-
1.83 /	1.82 /	1.83 /	-	1.83
1.81	-	1.81	-	-
1.68 /	-	1.68 /	-	-
1.55	-	-	-	-
-	1.34	1.34	-	-
1.06 ^	1.06 ^	1.06 ^	1.06 ^	1.06
0.98	-	-	-	-
0.54 /	0.54 /	0.54 /	-	0.54
-0.55 ^	-0.56 ^	-0.55 ^	-0.57 ^	-0.56
-	-0.63 /	-0.62 /	-0.64 /	-0.63
-	-	-	-0.74	-
-	-	-	-0.88	-
-	-	-0.99	-0.97	-
-	-	-	-1.01	-
-	-1.11	-	-1.12	-
-1.21 /	-1.22 /	-	-1.23 /	-1.22
-	-	-	-1.62	-

* relative retention times are used instead of absolute retention times. See section 4.3.1

" ^ " means all sources have the peak; " / " means over half the sources have the peak

Table 4.2. Extraction of "characteristic" peaks from the "effective" peaks. Sample: Baizhu

10 (or less) effective peaks in relative retention times, *t* *				Characteristic peaks
Sources of samples #				
Hong Kong	Beijing	Guangzhou	Reference	
-0.13 /	-0.13 /	-0.12 /	-	-0.13
-0.17 /	-0.17 /	-0.16 /	-	-0.17
-0.30	-	-	-	-
-0.46 /	-0.46 /	-0.46 /	-	-0.46
-0.60	-	-	-	-
-0.68 ^	-0.68 ^	-0.68 ^	-0.68 ^	-0.68
-0.72 ^	-0.74 ^	-0.73 ^	-0.73 ^	-0.73
-0.78 ^	-0.79 ^	-0.78 ^	-0.78 ^	-0.78
-0.82 /	-0.82 /	-0.82 /	-	-0.82
-	-	-	-1.19	-
-1.24 ^	-1.24 ^	-1.24 ^	-1.24 ^	-1.24
-	-	-	-1.27	-
-	-	-	-1.47	-
-	-	-1.55	-1.56	-
-	-	-2.25	-2.25	-
-	-	-	-2.73	-

* relative retention times are used instead of absolute retention times. See section 4.3.1

" ^ " means all sources have the peak; " / " means over half the sources have the peak

It should be noted that the number of "characteristic" peaks should be less than or equal to the number of "effective" peaks of individual herbal sample. The set of "characteristic" peaks can be changed when the herbal drug from a new source is added since the occurrence probability of the peaks may vary. However, the aforementioned rule still applies: when samples from more than half of the sources (over half of the occurrence probability) contain a particular peak for the herbal drug, the peak is considered to be one of the "characteristic" peaks. Hence, some peaks may be discarded or new candidates may be added. Thus, the "characteristic" peaks have to be amended when a new source is entered.

4.3 Library section

The library system or file of the database is actually an area for the collection of the qualitative and quantitative data for each herbal sample. The relative retention times, *t*, as well as the normalization factors of the effective peaks of each sample are calculated here.

4.3.1 Calculation of relative retention indices

The reproducibility of retention times depends on the stability of instrumental conditions such as carrier gas flow rate and temperature programming. It was found that such conditions of our GC/MS instrument was stable enough that the reproducibility of retention times is good. In some circumstances, however, the overall retention times of chromatographic peaks shifted together so that the absolute retention times may not be exactly the same in different injection trials. It may be due to the instrumental errors and the injection errors. In order to obtain comparable qualitative data, relative retention indices are used instead because they are not affected by overall shifting of retention times.

The relative retention times, *t*, of the "effective" peaks are calculated with reference to the retention time of internal standard peaks [12]. They are calculated by Equation 4.1.

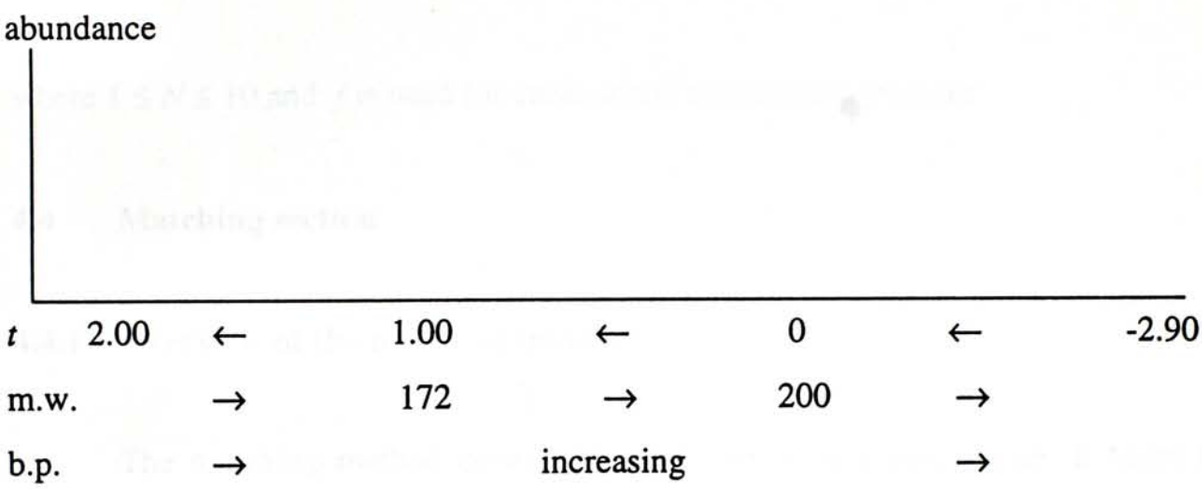
$$t = L_p / L_{IS} \qquad \text{(Eqn. 4.1)}$$

where L_p = retention time of IS (ethyl caprate) – retention time of the peak

L_{IS} = retention time of IS (ethyl caprate) – retention time of IS (ethyl caprylate)

The relative retention times show positive or negative values. Those with positive values mean that the peaks are eluted out faster than IS (ethyl caprate) while those with negative values are slower. The range of absolute retention times in the

"analysis window" is 300 to 1920 seconds while the range of relative retention times turn out to be 2.00 to –2.90 in descending order. Generally speaking, by referring to the position of the chromatographic peaks, the molecular weights or the boiling points of the components can be estimated, assuming the analytes interact on the same stationary phase of the column. The trend is shown in Figure 4.4.



t = relative retention times; m.w. = molecular weights; b.p. = boiling points

Fig.4.4. General trends of molecular weights and boiling points of components by referring to the relative retention times

Therefore, by referring to the absolute retention times or relative retention times of the components in some essential oils in the chromatographic patterns, their boiling points or molecular weights can be estimated.

The relative retention times of the 10 (or less) largest "effective" peaks of the herbal samples, which serve as known samples for library development, are tabulated in appendix D. The "characteristic" peaks of each herbal drug are also shown.

4.3.2 Normalization factors

Instead of using the absolute values, the peak area of each selected peak from the chromatogram was expressed as a percentage of the sum of the peak areas of all

selected peaks in each chromatogram [17]. This fraction is designated as the normalization factor. The normalization factor, f, of individual "effective" peak is calculated with the N largest effective peaks by Equation 4.2, with $1 \leq N \leq 10$. By using the sorting function, the largest peaks can be sorted out.

f = (area of effective peak among N peaks) / (total area of the N peaks) (Eqn.4.2)

where $1 \leq N \leq 10$ and f is used for comparison or matching purpose.

4.4 Matching section

4.4.1 Overview of the matching method

The matching method developed in this study is a simple one. It helps to assess the possibility of recognizing Chinese Medicinal Herbs using chromatographic information and a schematic methodology. The algorithms used here are simple calculations involving some logic functions. In this section, 10 or less most abundant effective peaks are used for comparison in most cases, because the error is smaller using major peaks. When the number of "effective" peaks or "characteristic" peaks of known herbal samples are less than 10 for the matching purpose, the remaining entries of relative retention times are assigned the value of 5.00 in the calculation area. On the other hand, the entries are assigned the value of –5.00 for the unknown in the input layout. The matching strategy compares the information of the unknown sample with those found in the library file. Matching of qualitative (relative retention times) and quantitative (normalization factors) data are carried out. On the basis of the matched result, the similarity score between the unknown sample and target herbal drug is calculated. A schematic flowchart showing the matching system is illustrated in Figure 4.5.

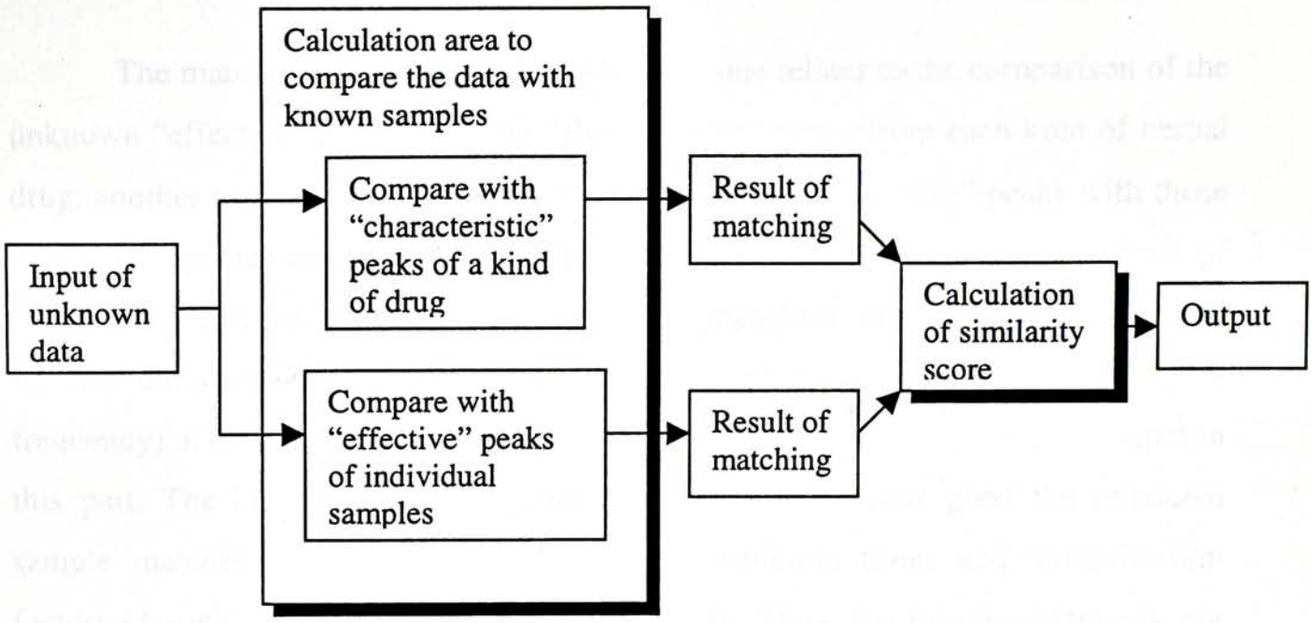


Fig. 4.5. A schematic flowchart showing the matching system

4.4.2 Input

The matching method consists of a spreadsheet which can be used to input the relative retention times and normalization factors of the unknown, which will then be compared automatically with those from the studied herbal samples in the file. The input layout is shown in Figure 4.6.

	P1*	P2	P3	P4	P5	P6	P7	P8	P9	P10
unknown										
	Information including relative retention times and normalization factors are input in this area									
abundance										

* "P" stands for peak

Figure 4.6. Input layout for unknown information

4.4.3 Matching strategy

The matching part involves two sections: one relates to the comparison of the unknown "effective" peaks with the "characteristic" peaks from each kind of herbal drug; another relates to the comparison of the unknown "effective" peaks with those from each species under study individually. From these two aspects, two kinds of information can be obtained. The former comparison provides information on whether the unknown sample has the "skeleton" peaks (which occurs with a higher frequency) of a particular kind of herb. Only relative retention times, are compared in this part. The latter comparison provide information on how good the unknown sample matches each candidate. The relative retention times and normalization factors of each effective peak are compared here. Thus, the matching strategy not only compares the whole picture (10 or less of the most abundant effective peaks) of the chromatographic patterns, but increases the weight or importance of the "characteristic" information.

4.4.4 Matching algorithms

The recognition of an unknown herbal sample is based on how good it matches the herbal drugs already studied. In this study, matching actually depends on some simple calculation procedure and logic functions. A similarity score is then given to each target herbal drug. The one with very high score is likely to be identical to the unknown sample, provided that the unknown is a member of the herbal drugs under study.

4.4.4.1 Matching with "characteristic" peaks

It was found that the maximum standard deviation of the relative retention times (analyte position) is 0.00577. At the confidence level larger than 90%, the significant separation of the chosen peaks should be equal or larger than 6 seconds. Thus, for the matching of "characteristic" peaks, each "effective" peak of the

unknown is compared with each "characteristic" peak from individual target herbal drug. The peaks are considered to be matched when the difference between their relative retention times are less than or equal to 0.02 (about 6 seconds), as shown in Equation 4.3.

$$|(t_i)_u - (t_j)_A| \leq 0.02 \qquad \text{(Eqn. 4.3)}$$

where $(t_i)_u$ is the relative retention time of the i th "effective" peak from the unknown sample; $(t_j)_A$ is the relative retention time of the j th "characteristic" peak from drug, A, for $i = 1, 2, \ldots, N_u$ for $N_u \leq 10$ where N_u is the number of "effective" peaks of the unknown; $j = 1, 2, \ldots, N_A$ for $N_A \leq 10$ where N_A is the number of "characteristic" peaks of drug A.

The total number of matched peaks is n_A and the maximum number of matching is N_A. For the matching of the unknown with the "characteristic" peaks from other herbal drugs, B, C, D, etc., the calculation is the same.

4.4.4.2 Matching with "effective" peaks

For the comparison of "effective" peaks from the unknown with those from each herbal species, Equation 4.3 is again used for comparing the relative retention times. The only difference is that each "effective" peak of the unknown is compared with each "effective" peak from individual species. The equation is replaced by Equation 4.4 for comparing "effective" peaks.

$$|(t_i)_u - (t_j)_a| \leq 0.02 \qquad \text{(Eqn. 4.4)}$$

where $(t_i)_u$ is the relative retention time of the i th "effective" peak from the unknown sample; $(t_j)_a$ is the relative retention time of the j th "effective" peak from

the species, a, for $i = 1, 2, \ldots, N_u$ for $N_u \leq 10$ where N_u is the number of "effective" peaks of the unknown; $j = 1, 2, \ldots, N_a$ for $N_a \leq 10$ where N_a is the number of "effective" peaks of the species, a.

The number of matched peaks is n_a and the maximum number of matching is N_a.

The normalization factors of the matched peaks are compared by calculating their ratios, m_j, using Equation 4.5.

$$m_j = \frac{\text{Min}((f_i)_u, (f_j)_a)}{\text{Max}((f_i)_u, (f_j)_a)} \qquad \text{(Eqn. 4.5)}$$

where $(f_i)_u$ is the normalization factor of the i th matched peak from the unknown

$(f_j)_a$ is the normalization factor of the j th matched peak from the herbal species, a.

If there are n_a matched "effective" peaks between the unknown and the herbal species a, the average matching ratios, M_a, of the normalization factors for these matched peaks is calculated by Equation 4.6.

$$M_a = \frac{\sum_{1}^{n_a} (m_j)}{n_a} \qquad \text{(Eqn. 4.6)}$$

4.4.5 Calculation of similarity scores

Assuming the highest score (identity match) is 100, the weighting distribution is shown in Figure 4.7. Since the abundance of components of the

essential oil composition can vary to some extent even within the same species, the normalization factor is given less weight (i.e. 20 out of 100) for the matching.

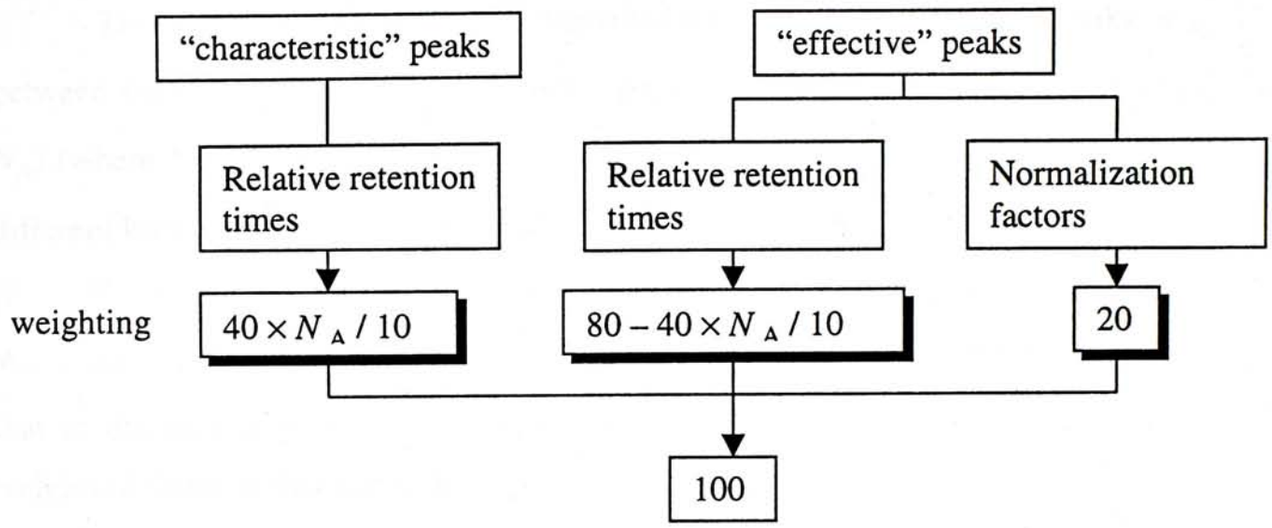


Figure 4.7. Weighting distribution in the scoring system

The similarity scores depend on the matching results described above. The scores are composed of three parts: result of matching relative retention times of "characteristic" peaks from different kinds of herbal drugs; results of matching relative retention times and normalization factors of the "effective" peaks from each species individually. It is noted from Figure 4.7 that the weighting for matching the "characteristic" peaks is higher when the number of peaks constituting the "characteristic" skeleton, N_A, is higher. It means that the importance of the matching with the "characteristic" peaks will be stressed. The total score, S_A, for drug A is calculated by using Equation 4.7.

$$S_A = [(n_A / N_A) \times (40 \times N_A / 10)] + [(n_{ai} / N_{ai}) \times (80 - 40 \times N_A / 10)] + M_{aj} \times 20 \quad \text{(Eqn. 4.7)}$$

where a_i is the species of drug A from source a_i, a_j is the species of drug A from source a_j, and a_i and a_j can be the same or different, depending on the matching results as explained below.

The first part of Equation 4.7 depends on the number of matched peaks, n_A, between the unknown and 'characteristic" peaks for herbal drug A. The ratio (n_A / N_A) (where N_A is the number of "characteristic" peaks for drug A) is taken because different kinds of herbal drugs have different number of "characteristic" peaks. ($40 \times N_A / 10$) is a weighting factor depending on the number of "characteristic" peaks. When the number, N_A, is small, the weighting factor in this part will be lower while that in the second part will be higher, or vice versa. Nonetheless, the maximum weighting factor in this part is 40.

The second part of Equation 4.8 depends on n_{ai}, which is the highest number of matched "effective" peaks between the unknown and that of herbal drug A from source a1 when the relative retention times are compared. Again, the ratio (n_{ai} / N_{ai}) (where N_{ai} is the number of "effective" peaks for drug A from source a_i) is taken for the same reason. ($80 - 40 \times N_A / 10$) is the weighting factor which depends on the number of "characteristic" peaks previously described. This part of the equation shows which source of drug A the unknown sample matches with higher similarity when the relative retention times are compared.

The third part of Equation 4.7 depends on the average matching ratio of the normalization factors, M_a, of the matched "effective" peaks of herbal drug A from a source and M_{aj} stands for the highest M_a values obtainable from drug A from the different sources. Sometimes, matching of relative retention times is good for a species but the matching of normalization factor is not, or vice versa. Thus, Equation 4.7 is so designed that n_a and M_a may be taken from different sources in order to

give the highest possible score. When matching for both the relative retention time and normalization factor are good for a species, then a_i is the same as a_j.

The calculations of similarity scores are repeated by matching the unknown with drugs B, C, D, ...to obtain S_B, S_C, S_D, ...etc.

4.4.6 Output

After calculating the similarity scores S_A, S_B, S_C, ... of herbal drugs A, B, C, ... respectively, sorting is carried out by using the sorting function of the Excel software, and a hit list of candidates can be obtained with their similarity scores is arranged in descending order. From the results, it can be deduced that the drug with the highest score may be the identity of the unknown sample, and hence the unknown sample may be "recognized".

Chapter 5: Performance of the proposed recognition system

5.1 Assessment of the performance of the proposed recognition scheme

5.1.1 Definition of similarity

A library search system facilitates the identification of an unknown from its chromatographic pattern by comparing it with each member of known patterns. In order to be useful for practical applications, a library system has to meet some criteria [30]. First, if the spectrum of the unknown at hand is in the library file, the system should be able to retrieve the respective spectrum. The process is defined as an "identity" search. If the unknown is not documented in the reference library, suitable reference (herbal drug) similar to the unknown should be retrieved. The process is defined as a "similarity" search. A recognition system requires a figure of merit to measure the similarity between the unknown and the reference. In the following section, a scoring scheme relating to similarity is discussed.

5.1.2 Performance test of the recognition method

5.1.2.1 Candidates in the library file

When the herbal drugs in the library file are treated as unknowns, they should be recognized and the expected herbal drug should have a high similarity score if the recognition system is effective. This is the "identity" search. In order to study this performance, the herbal drugs in the library file are treated as unknowns and their information is entered in the input layout. The similarity scores are calculated and the results are tabulated in Table 5.1. Codes are used here with explanation given at the end of the table.

Table 5.1. Searching results of unknown samples present in the library file (Performance test of the "identity" search)

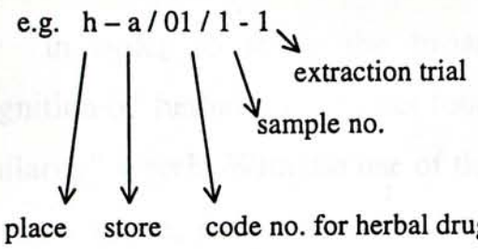
Drugs in file treated as unknowns #	Searching results for drug identity, the first three candidates in the hit list (score)			Searching performance *
h-a/01/1-1	01 (85)	23 (42)	15 (36)	+
h-d/01/1-1	01 (95)	23 (50)	02 (46)	+
h-e/01/1-1	01 (100)	15 (46)	23b (43)	+
h-a/02/1-2	02 (96)	09 (54)	15 (48)	+
h-d/02/1-4	02 (90)	04 (51)	15 (46)	+
h-e/02/1-2	02 (96)	01 (58)	09 (50)	+
b-a/02/1-1	02 (100)	09 (54)	15 (51)	+
h-a/04/1-1	04 (95)	18 (55)	25 (54)	+
b-a/04/1-1	04 (100)	18 (56)	15 (49)	+
g-a/04/1-1	04 (95)	18 (60)	02 (58)	+
h-a/05/1-1	05 (100)	06 (56)	20 (45)	+
b-a/05/1-1	05 (100)	06 (55)	20 (42)	+
g-a/05/1-1	05 (100)	06 (60)	15 (45)	+
h-b/06/1-1	06 (100)	05 (64)	20 (44)	+
h-d/06/1-1	06 (96)	29 (42)	11 (42)	+
g-a/06/1-1	06 (96)	05 (49)	20 (39)	+
h-a/07/1-3	07 (100)	09 (42)	25 (38)	+
h-d/07/1-1	07 (92)	08 (35)	09 (33)	+
b-a/07/1-2	07 (92)	12 (43)	01 (40)	+
g-a/07/1-2	07 (96)	23b (49)	15 (42)	+
f-b/08/1-1	08 (100)	25 (43)	09 (40)	+
h-a/09/2-3	09 (100)	09 (41)	08 (41)	+
h-e/09/1-1	09 (100)	25 (40)	19 (39)	+
g-a/09/1-1	09 (100)	25 (48)	02 (41)	+

h-b/11/1-1	11 (92)	12 (71)	07 (38)	+
h-d/11/1-1	11 (96)	12 (66)	06 (57)	+
h-e/11/1-1	11 (96)	12 (62)	23 (52)	+
b-c/11-b/1-1	11 (96)	12 (79)	09 (41)	+
h-d/12/1-1	12 (80)	11 (54)	06 (45)	+
b-a/12/1-1	12 (88)	11 (46)	09 (42)	+
g-a/12/1-1	12 (96)	11 (66)	02 (47)	+
b-c/12/1-1	12 (100)	11 (70)	29 (46)	+
f-b/12/1-1	12 (92)	11 (70)	29 (46)	+
h-a/15/1-2	15 (94)	04 (58)	02 (41)	+
h-d/15/1-2	15 (96)	02 (54)	04 (53)	+
h-e/15/1-1	15 (96)	09 (58)	04 (54)	+
h-f/15/1-1	15 (96)	02 (56)	04 (54)	+
h-b/18/1-1	18 (100)	04 (59)	25 (55)	+
h-d/18/1-1	18 (100)	25 (63)	04 (55)	+
h-a/19/1-2	19 (100)	09 (71)	18 (48)	+
h-d/19/1-2	19 (100)	18 (49)	23 (46)	+
h-e/19/1-1	19 (80)	02 (55)	15 (35)	+
h-d/20/1-3	20 (96)	11 (40)	05 (39)	+
h-e/20/1-1	20 (92)	05 (55)	06 (45)	+
f-a/20/1-2	20 (96)	23b (57)	23 (44)	+
g-a/20/1-2	20 (88)	29 (46)	23b (44)	+
h-a/23/1-5	23 (99)	23b (53)	20 (42)	+
h-d/23/1-4	23 (100)	23b (61)	09 (41)	+
h-e/23/1-1	23 (100)	23b (53)	29 (48)	+
h-f/23/1-1	23 (98)	23b (51)	29 (46)	+
b-a/23/1-2	23b (94)	20 (47)	23 (47)	+
g-a/23/1-1	23b(93)	23 (63)	20 (50)	+

h-a/24/1-1	24 (96)	08 (51)	18 (44)	+
h-d/24/1-1	24 (96)	18 (58)	08 (55)	+
b-a/24/1-1	24 (96)	09,18 (44)	08 (40)	+
h-a/25/1-1	25 (98)	08 (59)	18 (54)	+
h-d/25/1-1	25 (98)	18 (67)	09 (57)	+
b-b/25/1-1	25 (98)	08 (59)	18 (55)	+
h-a/29/1-1	29 (98)	23b (43)	20 (38)	+
f-a/29/1-1	29 (95)	23b (53)	23 (46)	+
b-b/29/1-1	29 (88)	23 (48)	12 (44)	+
h-a/30/1-1	30 (96)	23 (47)	05 (40)	+
b-a/30/1-1	30 (100)	23 (50)	06 (47)	+
f-b/30/1-1	30 (96)	23b (52)	06 (51)	+
std/01/1-1	01 (88)	23 (60)	15 (42)	+
std/02/1-1	02 (84)	20,23b (46)	19 (45)	+
std/04/2-1	04 (88)	18 (49)	15 (47)	+
std/05/1-1	05 (84)	06 (65)	20 (38)	+
std/06/1-1	06 (86)	20 (52)	11 (43)	+
std/06-b/1-1	06 (87)	20 (58)	05 (57)	+
std/07/1-1	07 (99)	01 (47)	09 (44)	+
std/08/1-1	08 (100)	25 (57)	09,11 (40)	+
std/09/1-1	09 (94)	02 (49)	04,19 (40)	+
std/11/1-1	11 (84)	12 (66)	02 (49)	+
std/12-b/1-1	12 (92)	11 (69)	07 (51)	+
std/15/1-1	15 (96)	23b (53)	01 (46)	+
std/18/1-1	18 (88)	19 (61)	23 (54)	+
std/20/1-1	20 (88)	06 (63)	11 (56)	+
std/23/1-1	23b (98)	23 (65)	29 (44)	+

* "+" means that the expected drugs with the highest score can be retrieved

Explanation of the codes is shown below:



Place: h: Hong Kong; b: Beijing
g: Guangzhou; f: Mongolia

store: a, b, c, etc. for different stores in that place

code no.: represents the herbal drug
e.g. 01 stands for Danggui
for other code nos., refer to Table 2.1

From the results calculated by the matching method, it is found that the average score for the candidates in the hit list with the highest score for the unknown sample is 95 with standard deviation 5. The score range for the first candidate in the hit list (estimated by 1 SD) is 90 – 100. The score ranges for the top candidates are summarized in Table 5.2.

Table 5.2. Scoring range for the first three candidates in the hit list for the search reported in Table 5.1

Priority	Average score	Standard deviation (SD)	Score range (estimate by 1 SD)
1	95	5	90 – 100
2	55	8	47 – 63
3	46	6	40 – 52

5.1.2.2 Unknown not found in the library file

In order to study the broader applicability of the matching method, the recognition of herbal samples not found in the library file should be assessed. This is the "similarity" search. With the use of the proposed experimental and instrumental analysis described above, their chromatographic patterns were obtained and interpreted by the aforementioned analysis procedure. New sets of data of relative retention times and normalization factors were extracted out. The data was input in the matching method and scores were calculated. The results are tabulated in Table 5.3.

Table 5.3. Searching results of unknown samples not present in the library file (Performance test of "similarity" search)

Unknown samples #	Searching results for drug identity, the first three candidates in the hit list (score)			Identity of the unknown with highest possibility *
h-b/01/1-1	01 (75)	23 (46)	18 (40)	Danggui +
h-d/02/1-1	02 (69)	15 (56)	12 (44)	Duhuo +
h-e/02/1-1	02 (68)	01 (47)	06 (43)	Duhuo +
std/04/1-1	04 (72)	11,15 (47)	25 (42)	Qianghuo +
h-d/05/1-1	05 (77)	06 (68)	20 (51)	Baizhu +
h-a/06/1-1	06 (58)	11 (41)	12 (40)	Canzhu +
h-a/07/1-1	07 (97)	18 (61)	09 (54)	Jingjie +
h-a/07/1-2	07 (96)	18 (54)	09 (53)	Jingjie +
b-a/07/1-1	07 (75)	18 (35)	12 (29)	Jingjie +
g-a/07/1-1	07 (75)	18 (51)	25 (32)	Jingjie +
h-d/09/1-3	09 (93)	07 (48)	02 (40)	Sharen +
h-a/11/1-1	12 (52)	11 (49)	29 (48)	Yujin -
h-a/12/1-1	11 (63)	12 (58)	02 (40)	Ezhu -
h-a/15/1-1	15 (83)	09 (58)	04 (55)	Chuanxiong +
h-d/15/1-1	15 (80)	23b (50)	02 (47)	Chuanxiong +
h-a/18/1-1	08 (71)	18 (70)	09 (52)	Qianhu -

h-a/19/1-1	19 (65)	02,04 (44)	09 (40)	Fangfeng +
h-a/20/1-2	20 (78)	12 (53)	11 (51)	Muxiang +
h-d/20/1-2	20 (64)	02 (44)	11 (43)	Muxiang +
f-a/20/1-1	20 (71)	05 (58)	06 (48)	Muxiang +
g-a/20/1-1	20 (77)	29 (58)	05 (56)	Muxiang +
h-a/23/1-2	23 (85)	23b (66)	01 (55)	Zisuye (purplish green) +
h-d/23/1-2	23 (96)	23b (64)	20 (42)	Zisuye (purplish green) +
b-a/23/1-1	23b (80)	23 (57)	20 (44)	Zisuye (greenish) +
g-a/25/1-1	25 (82)	08 (61)	04,12 (44)	Gaoliangjiang +
g-a/29/1-1	29 (86)	23b (48)	23 (46)	Peilan +
g-a/30/1-1	30 (84)	23b (45)	06,23 (43)	Huoxiang +

* "+" means correct match can be achieved; "–" means incorrect match is found

Explanation of the codes is the same as for Table 5.1

From results in Table 5.3, it is found that the average score for the top candidate in the hit list was 76 with a standard deviation of 12 and the range of the score (estimated by 1 SD) is 64 – 88. The score range for the top candidates in the hit lists is summarized in Table 5.4.

Table 5.4. Scoring range for the first three candidates in the hit list for the search reported in Table 5.3

Priority	Average score	Standard deviation (SD)	Score range (estimated by 1 SD)
1	76	12	63 – 88
2	53	9	44 – 62
3	45	7	38 – 52

5.1.3 Information drawn from the scores

5.1.3.1 Recognition of the unknown sample in terms of similarity

The score range calculated in the "identity" search shows that the first candidate in the hit list has a higher score than that of the second one. The difference between their ranges (90 – 100 and 47 – 63 respectively) is large. Thus, clear differentiation of the matching priorities can be drawn and the "identity" matching using the proposed scoring method is sufficiently good. However, for the second and the third candidates, ambiguity is found in the lower score region of the second one unless a higher score can be reached.

For the "similarity" search, it is found that the score range of the top candidate in the hit list is lower when compared with that for the "identity" search. This range seems to fill the gap between the first and second candidates in the "identity" search. Besides, for the "similarity" search, the difference between the first and the second candidates are not so clear cut as those found in the "identity" search. It is expected because the unknown candidates are not part of the library file, from where a totally matched score can be achieved. Therefore, the term "similarity" is used in real situation where the unknown is most likely not contained in the library file. Though the identities of the unknown samples cannot be a hundred percent ascertained, they can be recognized in terms of the degree of "similarity". A diagram for clarification is shown in Figure 5.1.

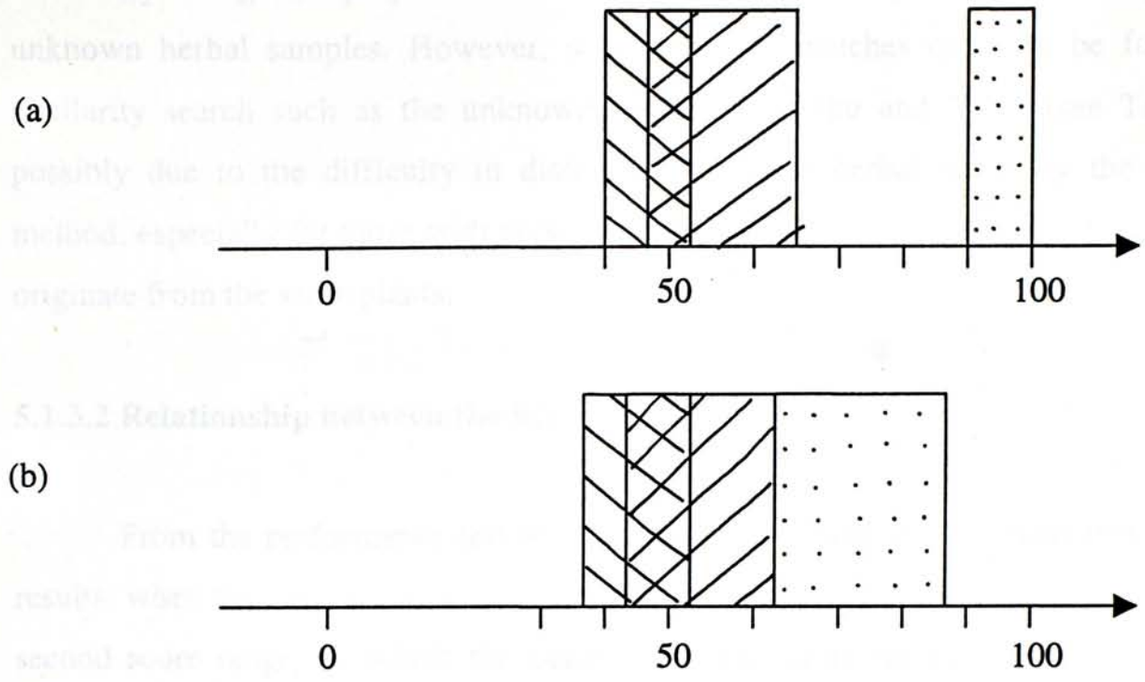


Figure 5.1. (a) Score range of "identity" search; (b) score range of "similarity" search

There are several points concerning the scores which may facilitate the recognition of an unknown herbal drug which belongs to a kind of the herbal drugs under study.

(a) When the scores of the first and second candidates in the hit list fall in different score range as shown in Figure 5.1, the candidate is very likely to be the identity of the unknown sample.

(b) When their scores fall in the same high score range, the matching will depend on the extent of difference between them.

- If their difference is still large, the one with the higher score is more likely to be the identity of the unknown sample.
- If their difference is small, both candidates are likely to be the identity of the unknown sample.

(c) When their scores fall in the same low score range, both candidates are less likely or unlikely to be the identity of the unknown sample, depending on the scores.

By using the proposed methodology, it is often possible to recognize the unknown herbal samples. However, some negative matches can also be found in a similarity search such as the unknown samples of Ezhu and Yujin (see Table 5.3), possibly due to the difficulty in distinguishing some herbal drugs by the proposed method, especially for those with very close relationships such as Ezhu and Yujin which originate from the same plants.

5.1.3.2 Relationship between the herbal drugs

From the performance test of the matching method, it was found that for some results, when the first and second matches are both in the high score range (higher than second score range) or within the same range with close scores, the first and second matches are somehow related due to the similarity of their chromatographic patterns or the composition of their essential oils. Some examples are illustrated in the following points:

(a) It is noted that herbal drugs Baizhu (05) and Canzhu (06) (see both Tables 5.1 and 5.3) usually come together as the top two candidates in the hit list. From the Chinese Pharmacopoeia [31], the scientific names for the two herbal drugs are Atractylodes mcrocephala Koidz. and Atractylodes lancea (Thunb.) DC. respectively. They come from the same genus of Atractylodes and their therapeutic functions are similar. Thus, the scores between them are close.

(b) Close scores are usually found for herbal drugs Ezhu (11) and Yujin (12). Sometimes, unknown samples are wrongly matched. These two herbal drugs are rather difficult to be distinguished by only comparing their chromatographic patterns of essential oils. The composition of their essential oils are rather similar. Actually, Ezhu and Yujin originate from the same genus and sometimes even from the same plants but in different parts, and their therapeutic functions are similar. It may be the reason for the high similarity of their chromatographic patterns and thus, close scores can be obtained.

5.2 Applicability of the proposed methodology

Overall, with the proposed methodology including the experimental and instrumental analysis of essential oils, and the recognition scheme, identification of herbal drugs with the use of chromatographic patterns is possible. Here, herbal drugs can be recognized in a broader sense i.e. between genera and among species with the same genus. The unknown drug is identified with the common name used in Chinese Pharmacopoeia (or retail name) instead of the accurate botanical name. Thus, the recognition can be used as a pre-recognition process. Besides, it is useful when the herbal samples are in crushed or powder form which may cause difficulty for expert examination.

Generally speaking, the more specific components an oil contains, the easier the herbal sample can be recognized. The ease in distinguishing the chromatographic pattern decreases in the following order: (a) oils containing many specific components at significant levels; (b) oils containing a single major component constituting 70% or more of the total composition, and (c) oils consisting mostly of terpene hydrocarbons [14]. The chromatographic patterns of the essential oils give us insight about the oil composition. Together with the MS information, some components can be identified.

From the aforementioned study, the scores obtained for the proposed matching method may throw some light on the relationship between herbal drugs from a scientific point of view.

Furthermore, larger range of herbal samples with great variation of essential oil abundance can be studied by using the proposed standardized procedure, which can therefore be used to analyze a wide variety of herbal drugs.

5.3 Limitation of the proposed methodology

As mentioned before, the herbal samples is recognized in a broad sense, refined differentiation is difficult using the simple calculation.

As different locations, different sources, time and method of cultivation, season, etc. cause changes in the aroma composition, those herbs with very close relationships such as Ezhu and Yujin impose ambiguity in recognizing the unknown samples of such kind of herbs solely by comparing the chromatographic patterns of their essential oils.

A successful identification of an essential oil largely depends on the correct experimental conditions, the very nature of each oil, and the complexity of the sample. Each laboratory would have to build its own library file because the reproducibility of retention time and mass ion intensity often vary much depending on experimental conditions, such as column condition and ion source tuning. Library files generated in one laboratory might not be compatible with the experimental conditions set in another laboratory [14]. Thus, the experimental and instrumental conditions as well as the analysis procedure need to be exactly followed.

5.4 Future prospect

The methodology built up in this research can be further developed.

A systematic method can be developed for the recognition of Chinese Medicinal Herbs. Through studying the chromatographic patterns and the mass spectra of the of the essential oil, valuable information can be obtained. Besides, the relationship between the herbal samples can be studied from the chemical point of view. As the present methodology deals with the recognition of the herbal samples in a broad sense described above, finer differentiation is not considered by adopting a simple calculation method. It

may be possible to perform further differentiation if more comprehensive species with accurate botanical identities can be obtained for library development.

By studying more samples, the size of the database can be increased and thus the applicability of the methodology is also increased. This systematic scheme has the merits of being less time-consuming and requiring small amounts of sample, as well as requiring no expert knowledge on Chinese Medicinal Herbs.

Apart from the chromatographic retention indices, mass spectral information can also be obtained from GC/MS analysis. The mass data of individual components provide auxiliary information. From the review study, it was reported that dried plant material was analyzed by directly introducing it into a mass spectrometer [4]. From this idea, instead of acquiring each spectrum for the corresponding component in a mixture, one "complete" mass spectrum for the mixture may be obtained by using the EI mode technique. By studying the mass spectral pattern in this manner, the "decisive" peaks can be "extracted" by referring to the unique mass-to-charge ratios of the ions, which can provide the structural and mass information of the sample mixture.

Chapter 6: Conclusion

In this study, a methodology has been proposed for the recognition of Chinese Medicinal Herbs, using the gas chromatographic patterns of essential oils extracted from the herbal samples. In order to achieve this, the experimental and instrumental conditions, as well as the analysis procedure, are standardized in order that reasonable comparison can be made. The process of analysis and the results have been discussed in the previous chapters.

The Dean and Stark-type trap apparatus is used for the extraction of essential oils and is simple. Less samples can be used and hence further enrichment of the extracted oils is not required. It was found that the major components extracted by the proposed method were usually comparable with those found in the corresponding herbal drugs as reported in literature. Although the chromatographic patterns may not reveal the actual content of the components present in the herbal drugs, the reproducibility of the patterns is satisfactory in qualitative aspect. On the other hand, the average relative deviation (including injection errors) in quantitative analysis is about 15%. In order to increase the reliability of the methodology, peaks with higher abundance are chosen for further analysis. The experimental section of the methodology has been discussed in Chapter 2.

Modern capillary gas chromatography show high resolving power even for complex mixtures. Chromatographic patterns are specific and they can be used as a fingerprint of herbal drugs containing essential oils. The precision of our instrument is good with average relative standard deviation of about 2%. GC is used to preliminary observe the composition of the extracted oil, including the relative abundance of the components. Due to the great variation of the oil contents between different species, dilution strategy is proposed to prevent overloading of the column and provide suitable "analysis window" when using GC/MS. The extent of dilution depends on the R values

(Eqn. 3.1). From this, the chromatograms can be better used for comparison with the concentrations of the components controlled below about 100 ppm with the most abundant peak from the component with at least 40 ppm.

With the combination of mass spectrometry, GC/MS is a powerful tool for identification of the trace components. The value of mass spectra lies in their high information content. MS data serve as an auxiliary tool to distinguish chromatographic peaks for comparison, even for the peaks which are close together. Through analysis of relative retention times and normalization factors, the chromatograms obtained by GC/MS are used for database development. The precision of our GC/MS instrument is good with average relative standard deviation of about 4%. The concentration range under study is linear. The instrumental analysis has been discussed in Chapter 3.

For the recognition of the Chinese Medicinal Herbs, a scheme consisting of the chromatographic analysis and matching methods has been developed. "Effective" peaks are extracted out from the chromatographic patterns of GC/MS by some proposed criteria based on the relative abundance of the components and the threshold values for integration of peak areas. Relative retention indices are assigned to the "effective" peaks through calculation using Eqn. 4.1. The added internal standards are mainly used for this calculation. These relative retention indices serve as codes for matching purpose. Besides, the normalization factors of the "effective" peaks are calculated by Eqn. 4.2 instead of absolute abundance. These normalization factors are also used for comparison. Apart from the "effective" peaks, "characteristic" peaks are chosen for corresponding herbal drugs to facilitate recognition.

On the basis of using "effective" peaks, "characteristic" peaks and normalization factors, matching algorithms have been designed for the calculation of similarity score so that the unknown samples can be identified with a certain confidence. It has been found that promising results of recognition of herbal samples can be obtained through

the performance test. Thus, recognition of unknown samples is possible. Moreover, some herbal samples with similar matching results are due to the close relationship, such as same genus and similar therapeutic indices between the species, from the chemical point of view. The methodology deals with the chromatographic analysis and recognition matching scheme has been discussed in Chapter 4. The performance test for the methodology has been investigated in Chapter 5.

In conclusion, a methodology has been introduced, which offers a simple, user-friendly and systematic way for the recognition of Chinese Medicinal Herbs containing essential oils by gas chromatography. It allows us to recognize the herbal sample, requiring no special professional knowledge on CMH. Nonetheless, the methodology still has its own limitation which needs further improvement in future work. The applicability and the limitation of the methodology has been discussed in Chapter 5. Besides, the mass spectral data obtained from GC/MS analysis can provide a valuable supportive information for the potential use of direct MS patterns for recognition.

References

1. 中草藥有效成分提取與分離，中國科學院藥物研究所，上海科學技術書出版社.

2. Huang, K. C. *The Pharmacology of Chinese Herbs*, CRC Press Inc.: Florida, 1993.

3. Liu, C. X.; Xiao, P. G. *An Introduction to Chinese Materia Medica*, Peking Union Medical College Beijing Medical University Press: Beijing, 1993.

4. Brunke, E. J. *Progress in Essential Oil Research: Proceedings of the International Symposium on Essential Oils*, W. de Gruyter: W. de Gruyter; New York, 1986.

5. 藥物分析雜誌， 1987 ， 7(2)， 117.

6. Teranishi, R.; Buttery, Ron G.; Sugisawa, H. *Bioactive Volatile Compounds from Plants*, American Chem. Soc.: Washington, D. C., 1993.

7. Reid, Daniel P. *A Handbook of Chinese Healing Herbs*, Shambhala Pub. Inc.: Boston, 1995.

8. Rijks, J.; Cuvers, J.; Noy, Th.; Cramers, C. *J. Chromatogr.* **1983**, *279*, 395.

9. Shan, L. D. *Zhongyao Jiandinxue*, Beijing Yikedaxue & Beijing Xiehe Yikedaxue Lianhe Chubanche: Beijing, 1991.

10. Lu, Q. F.; Shun, X. Q.; Ben, C. N.; Lu, H. Q.; Gao, W. *Beijing Zhongyiyaodaxuexuebao*, **1995**, *18 (3)*, 64.

11. Zhang, L. H.; Wang, H. Y.; Liu, J. *Shengyang Yaoxuyuan Xubao*, **1995**, *11*, 112.

12. Tian, J. G.; Xu, X. R. *Zhongchaoyao*, **1989**, *20*, 221.

13. Sandra, P.; Bicchi, C. *Capillary Gas Chromatography in Essential Oil Analysis*, Heuthig: Heidelberg; New York, 1987.

14. Chien, M. *Anal. Chem.* **1985**, *57*, 348.

15. Park, M. K.; Cho, J. H.; Kim, N. Y.; Park, J. H. *Anal. Chim. Acta* **1993**, *284*, 73.

16. Bicchi, C.; D'Amato, A.; Nanao, G. M.; Frattini, C. *J. Chromatogr.* **1983**, *279*, 409.

17. Godefroot, M.; Sandra, P.; Verzele, M. *J. Chromatogr.* **1981**, *203*, 325.

18. Richmond, R.; Pombo-Villar, E. *J. Chromatogr. A* **1997**, *760*, 303.

19. Linskens, H. F.; Jackson, J. F. *Modern Methods of Plant Analysis: Plant Volatile Analysis*, Springer: Berlin; New York, 1997.

20. Jayatilaka, A.; Poole, S. K.; Poole, C. F.; Chichila, Tina M. P. *Anal. Chim. Acta* **1995**, *302*, 147.

21. Marotti, M.; Piccaglia, R.; Giovanelli, E. *J. Agric. Food Chem.* **1996**, *44*, 3926.

22. 中草藥學，四川醫學院，人民衛生出版社， 1982.

23. Chen, Y.; Li, Z.; Xue, D.; Qi, L. *Anal. Chem.* **1987**, *59*, 744.

24. Baugh, P. J. *Gas Chromatography: A Practical Approach*, IRL Press at Oxford University Press: Oxford; New York, 1993.

25. Jennings, W. G. *Analytical Gas Chromatography*, Academic Press Inc.: Orlando; Florida, 1987.

26. Miller, J. C.; Miller, J. N. *Statistics for Analytical Chemistry*, Ellis Horwood Ltd.: New York, 1993.

27. Finnigan Mat Magnum GC/MS Operators Manual, 1992.

28. Finnigan Mat Magnum GC/MS Software Manual, 1992.

29. Long, Gary L.; Winefordner, J. D. *Anal. Chem.*, **1983**, *55 (7)*, 712A.

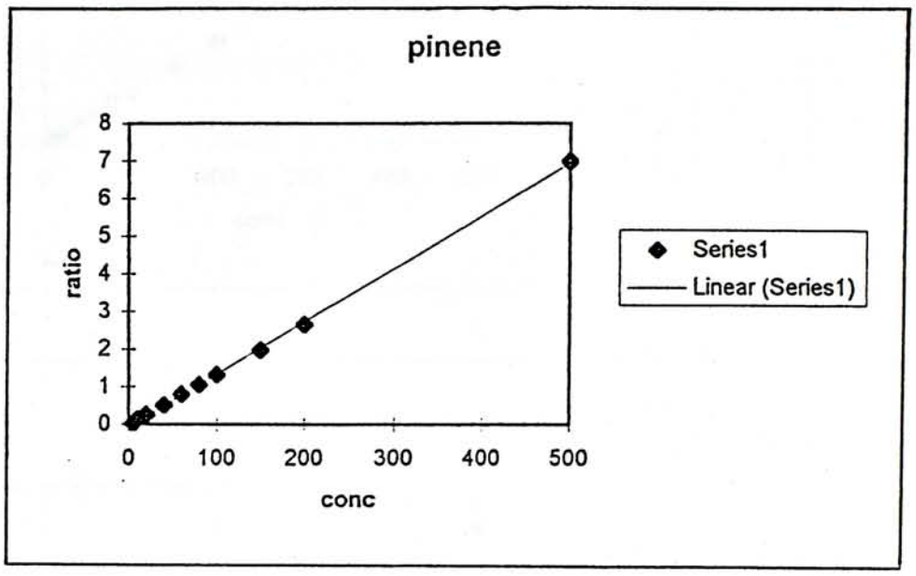
30. Ruprecht, M.; Clerc, J. T. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 241.

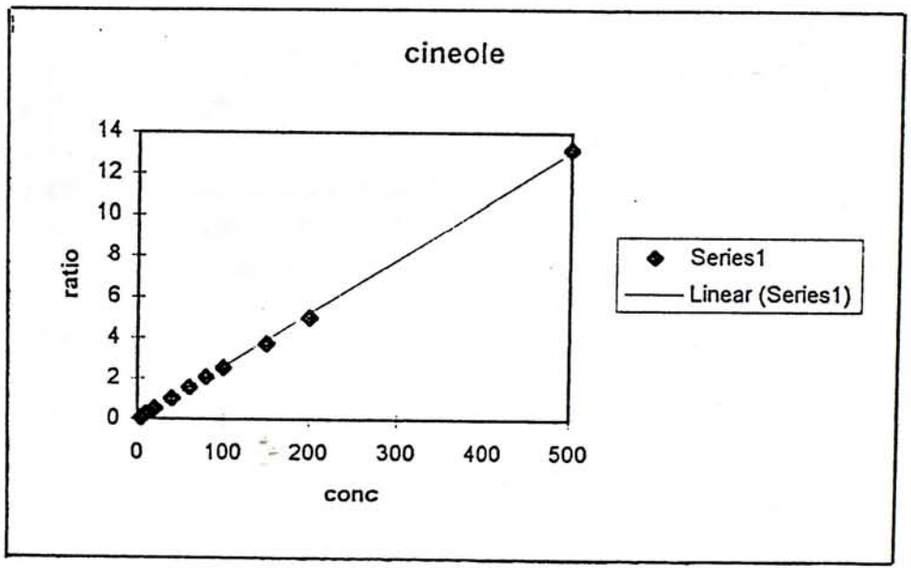
31. 中華人民共和國藥典

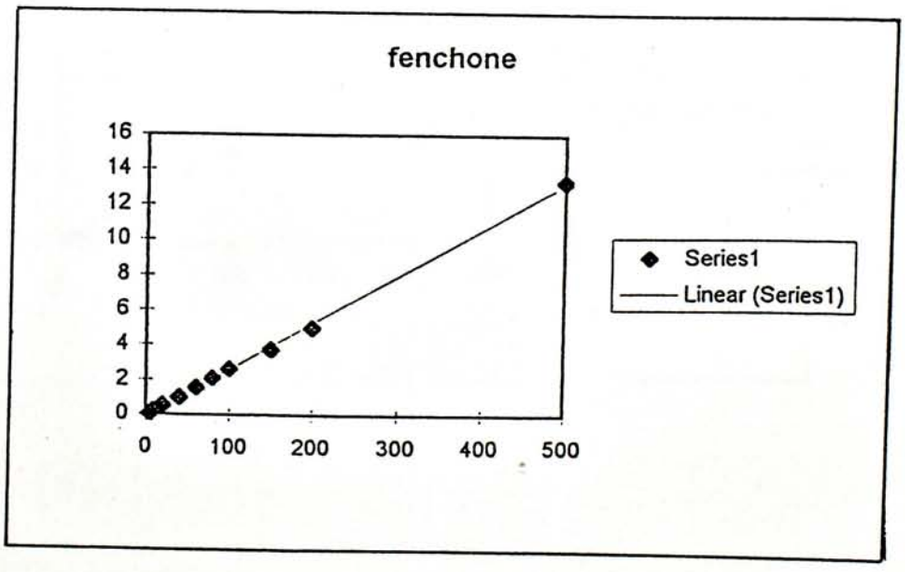
Appendix

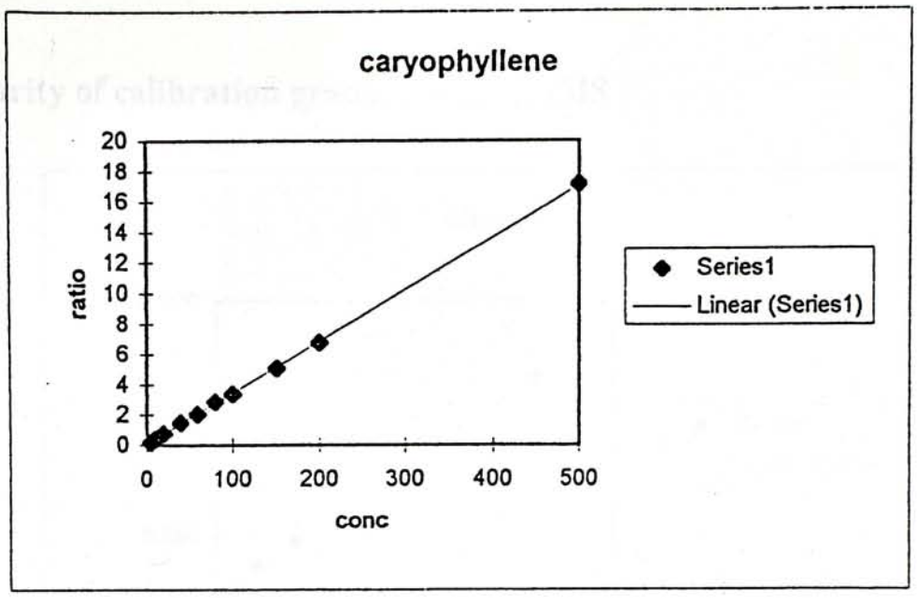
A1

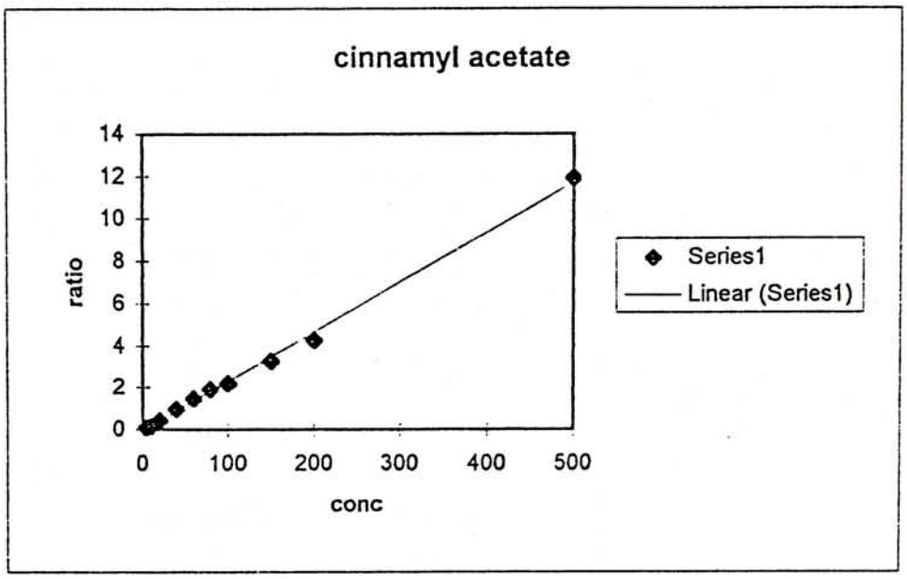
A. Linearity of calibration graphs using GC

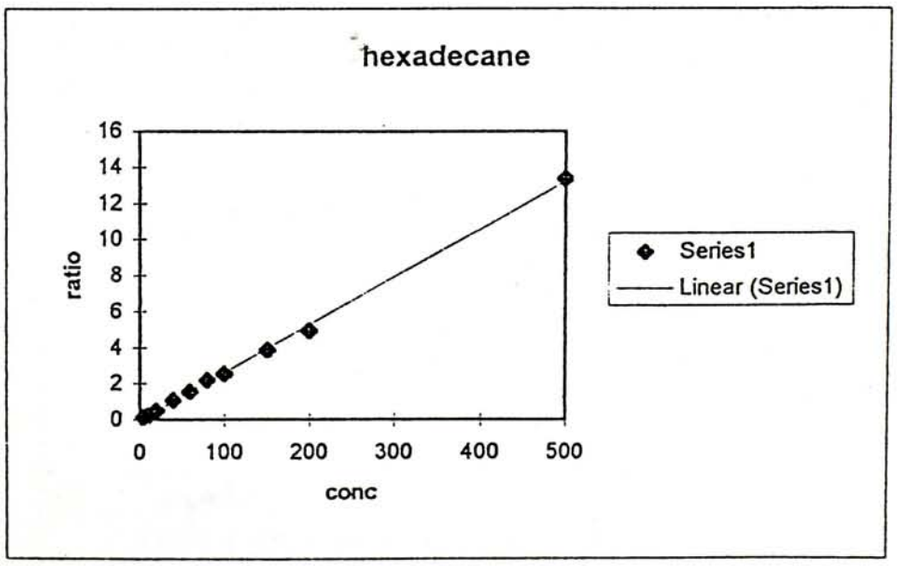






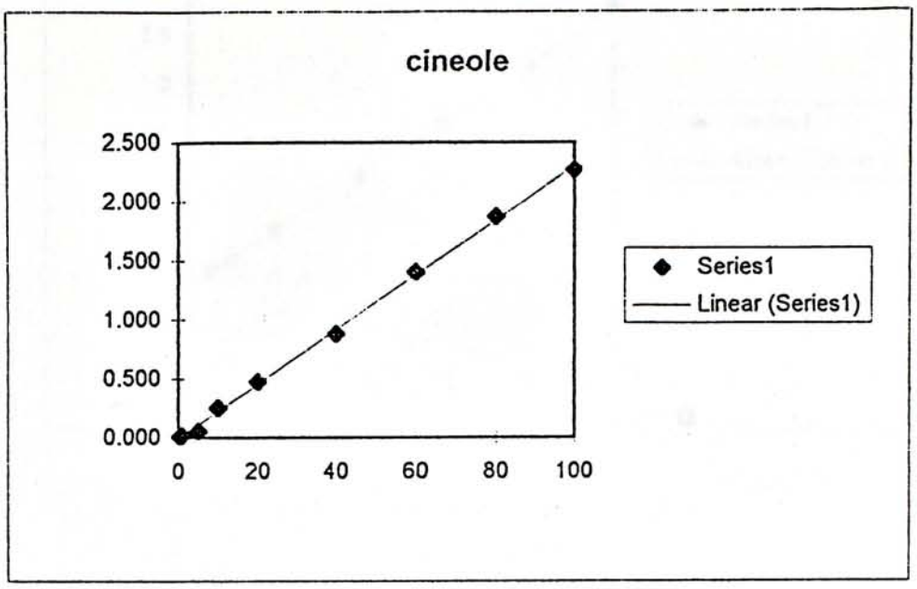

caryophyllene
ratio
conc
0
2
4
6
8
10
12
14
16
18
20
0
100
200
300
400
500
Series1
Linear (Series1)

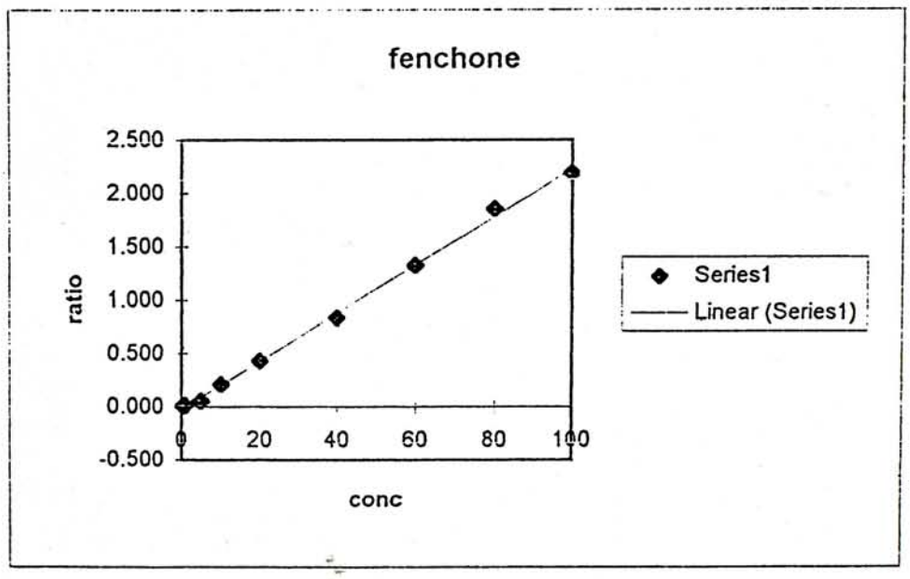

cinnamyl acetate
ratio
conc
0
2
4
6
8
10
12
14
0
100
200
300
400
500
Series1
Linear (Series1)

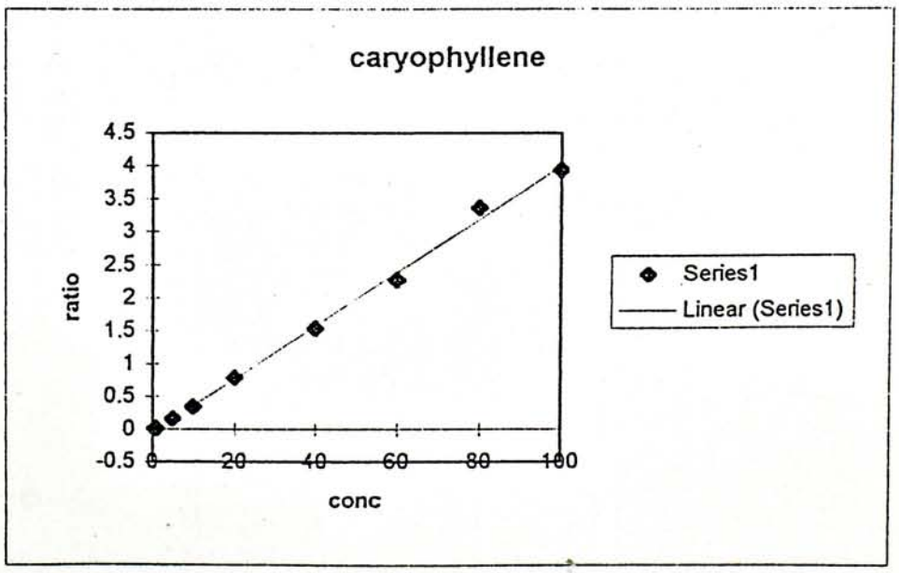

hexadecane
ratio
conc
0
2
4
6
8
10
12
14
16
0
100
200
300
400
500
Series1
Linear (Series1)

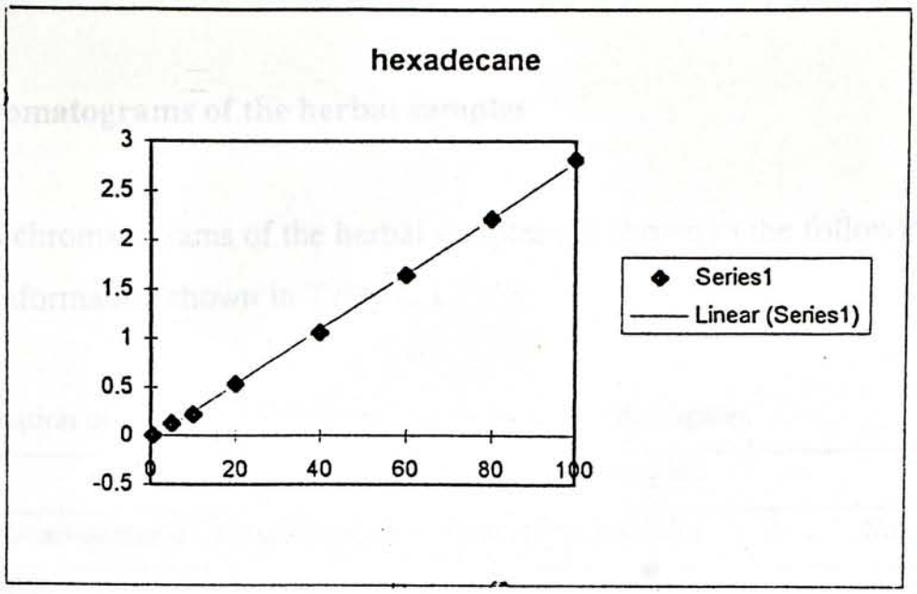
B. Linearity of calibration graphs using GC/MS

A3








hexadecane
3
2.5
2
1.5
1
0.5
0
-0.5
0
20
40
60
80
100
Series1
Linear (Series1)

C. GC/MS chromatograms of the herbal samples

The GC/MS chromatograms of the herbal samples are shown in the following Figure C.1, with their information shown in Table C.1.

Table C.1. Information of the GC/MS chromatograms shown in the figures

Herbal medicines treated as known samples		
Figure	Retail names or scientific names of the herbal samples	Sources*
1	Danggui	h-a
2	Danggui	h-d
3	Danggui	h-e
4	Duhuo	h-a
5	Duhuo	h-d
6	Duhuo	h-e
7	Duhuo	b-a
8	Qianghuo	h-a
9	Qianghuo	b-a
10	Qianghuo	g-a
11	Baizhu	h-a
12	Baizhu	b-a
13	Baizhu	g-a
14	Canzhu	h-b
15	Canzhu	h-d
16	Canzhu	g-a
17	Jingjie	h-a
18	Jingjie	h-d
19	Jingjie	b-a
20	Jingjie	g-a
21	Bajiao	f-b
22	Sharen	h-a
23	Sharen	h-e
24	Sharen	g-a
25	Ezhu	h-b

26	Ezhu	h-d
27	Ezhu	h-e
28	Wenezhu (Curcuma wenyujin)	b-c
29	Yujin	h-d
30	Yujin	b-a
31	Yujin	g-a
32	Pianjianghuang (Curcuma wenyujin Y. H. Chen et Ling) (Rhizoma Wenyujin concisa)	b-c
33	Wenyujin (Curcuma wenyujin)	f-b
34	Chuanxiong	h-a
35	Chuanxiong	h-d
36	Chuanxiong	h-e
37	Chuanxiong	h-f
38	Qianhu	h-b
39	Qianhu	h-d
40	Fangfeng	h-a
41	Fangfeng	h-d
42	Fangfeng	h-e
43	Muxiang	h-d
44	Muxiang	h-e
45	Muxiang	f-a
46	Muxiang	g-a
47	Zisuye (purplish green)	h-a
48	Zisuye (purplish green)	h-d
49	Zisuye (purplish green)	h-e
50	Zisuye (purplish green)	h-f
51	Zisuye (greenish)	b-a
52	Zisuye (greenish)	g-a
53	Xiangru	h-a
54	Xiangru	h-d
55	Xiangru	b-a
56	Gaoliangjiang	h-a
57	Gaoliangjiang	h-d
58	Gaoliangjiang	b-b

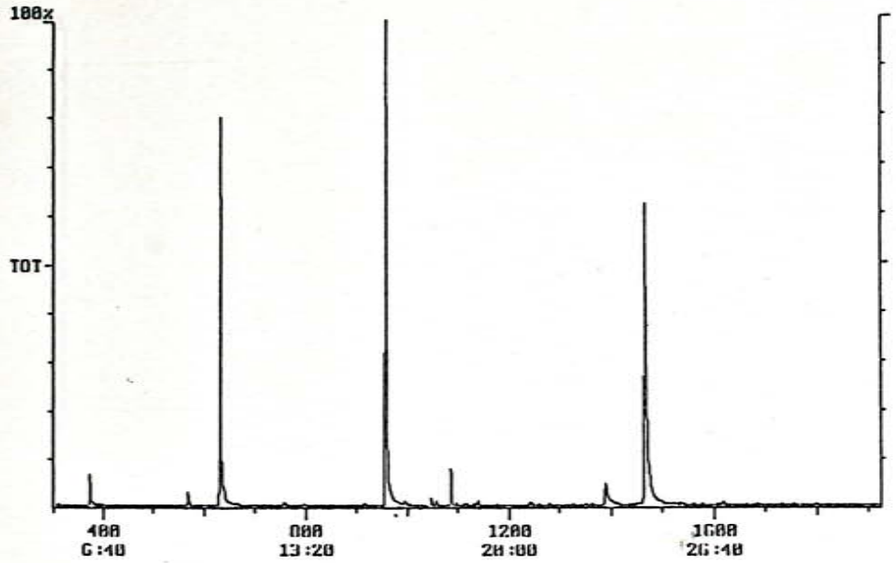
59	Peilan	h-a
60	Peilan	f-a
61	Peilan	b-b
62	Huoxiang	h-a
63	Huoxiang	b-a
64	Huoxiang	f-b

Reference herbal samples

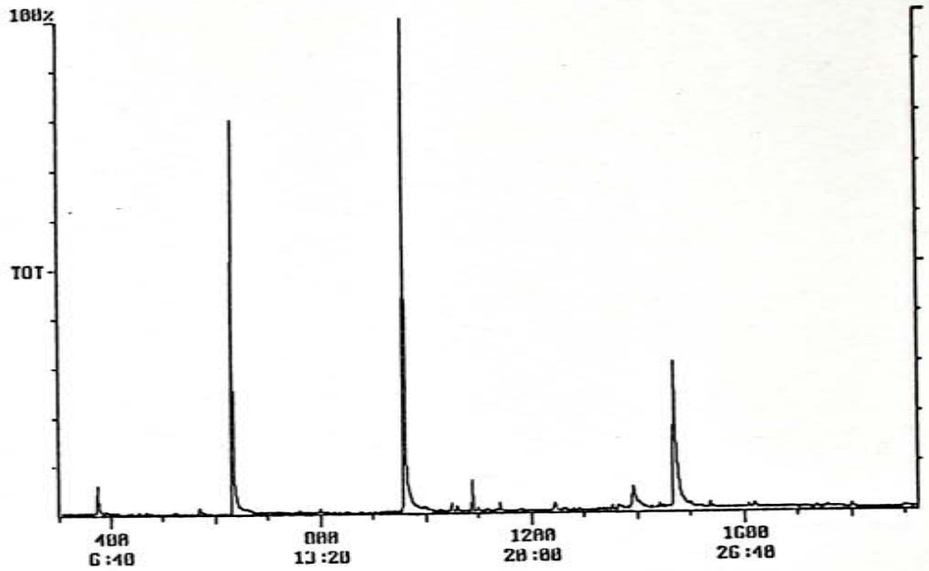
Figure	Retail names or scientific names of the herbal samples	Sources
65	Angelica sinensis (Oliv.) Diels	std
66	Angelica pubescens Maxim. F. biserrata Shan et Yuan	std
67	Notopterygium incisum Ting ex H. T. Chang	std
68	Atratylodes macrocephala Koidz	std
69	Atratylodes lancea (Thumb.) DC.	std
70	Atratylodes chinensis (DC.) Koidz	std
71	Schizonepeta tenuifolia Briq.	std
72	Illicium verum Hook.f.	std
73	Amomum villosum Lour.	std
74	Curcuma phaeoculis Valeton	std
75	Ligusticum chuanxiong Hort.	std
76	Peucedanum praeruptorum Dunn	std
77	Ledebouriella divaricata (Turcz.) Hiroe	std
78	Aucklandia lappa Decne.	std
79	Perilla frutescens (L.) Britt.	std

*Note: h-a means sample from Hong Kong, store A; h-b means samples from Hong Kong, store B, etc.; b-a means sample from Beijing, store A, etc.; g-a means sample from Guangzhou, store A, etc.; f-a means sample from Mongolia, store A, etc.; std means reference samples from store selling authentic samples.

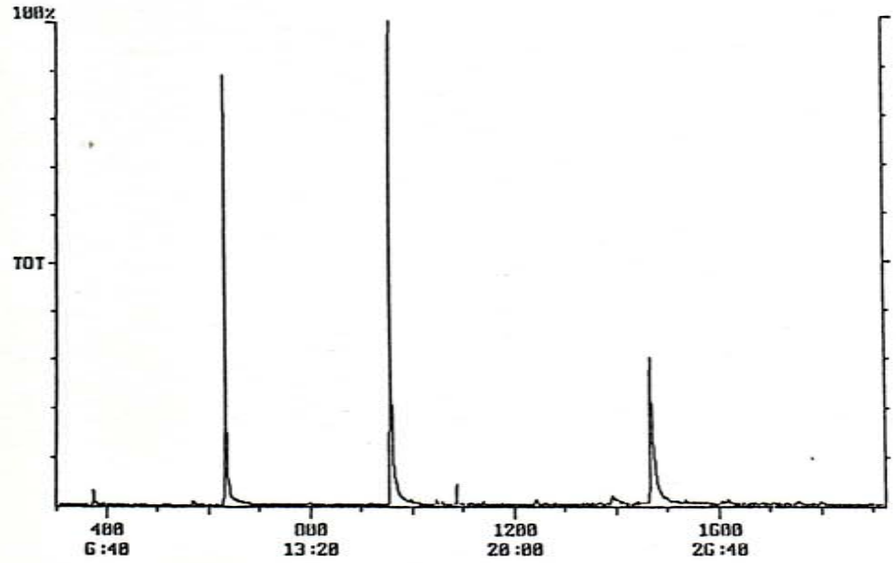
1


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

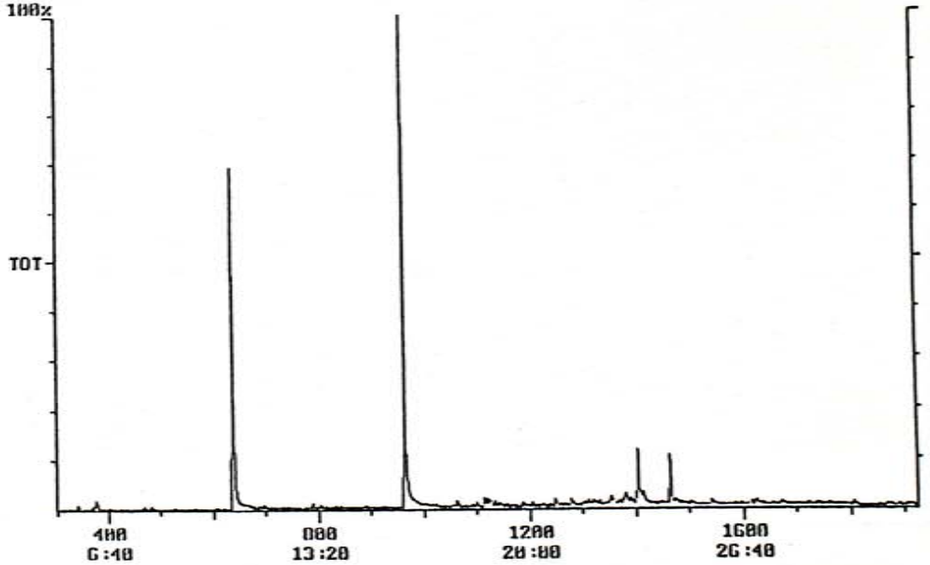
3


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

2

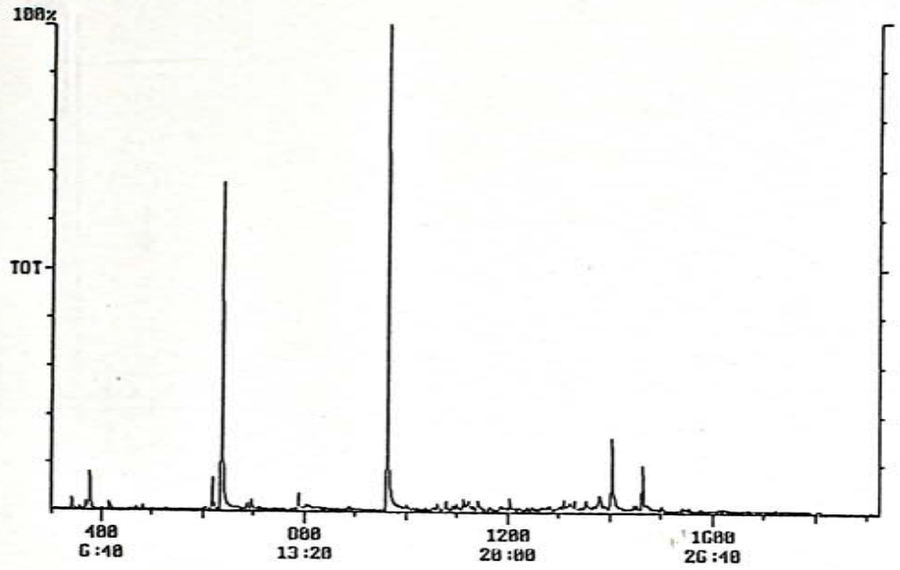

100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

4

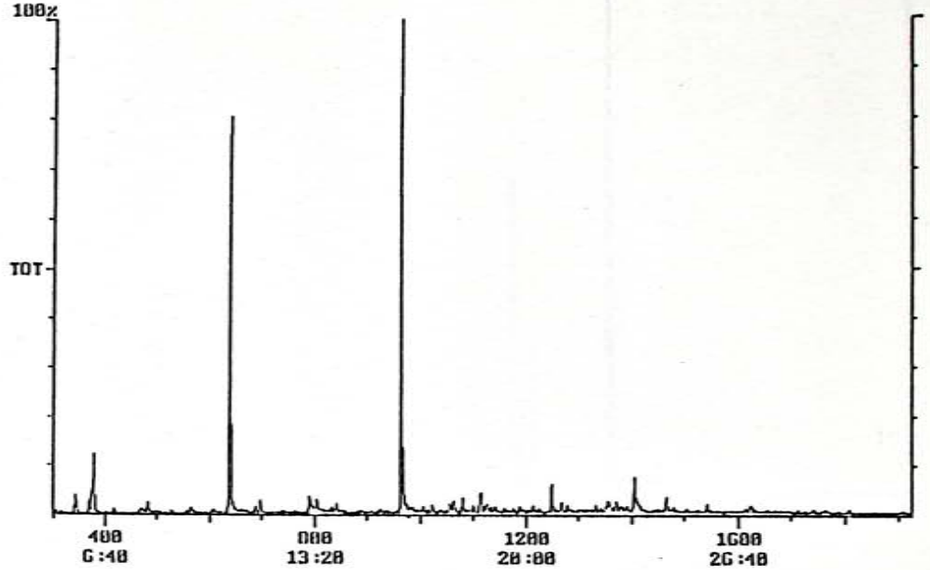

100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

48

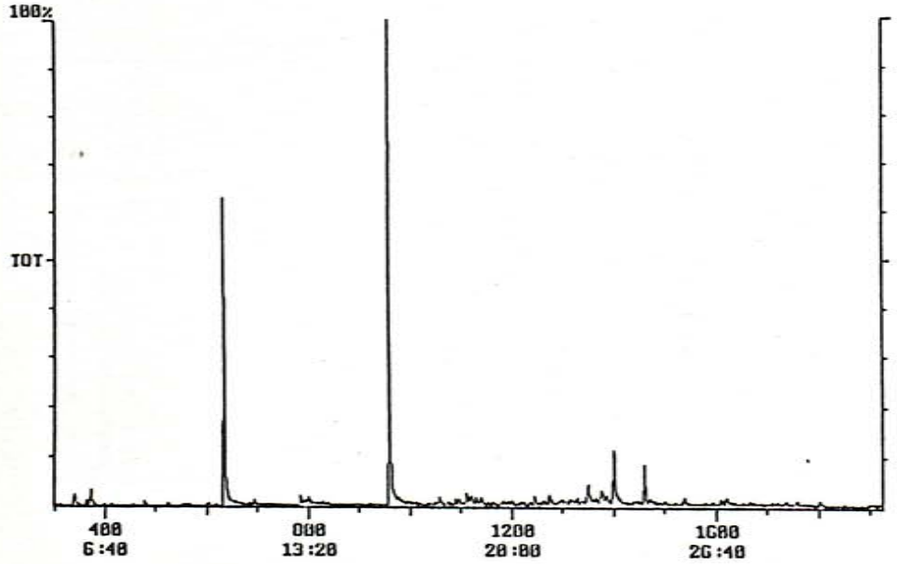
5


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

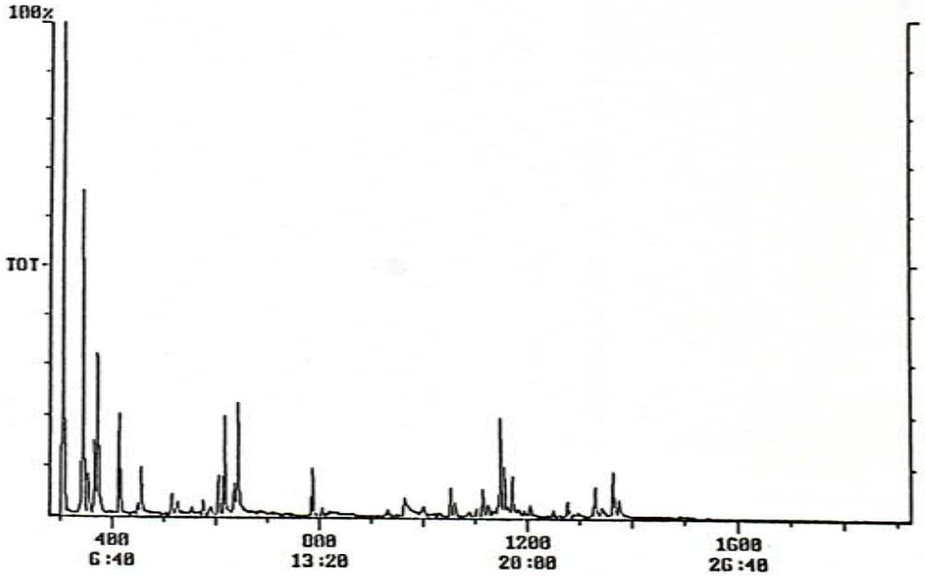
7


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

6

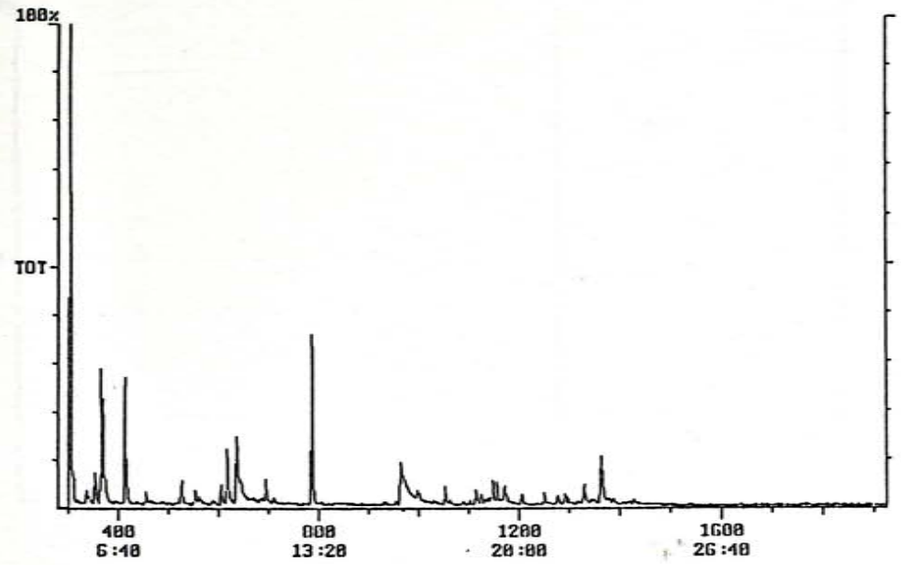

100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

8

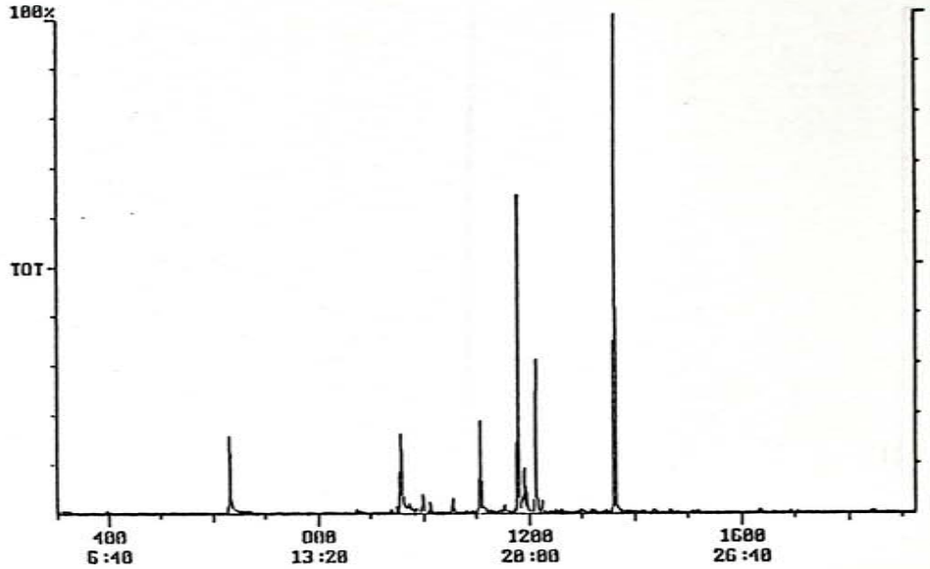

100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

49

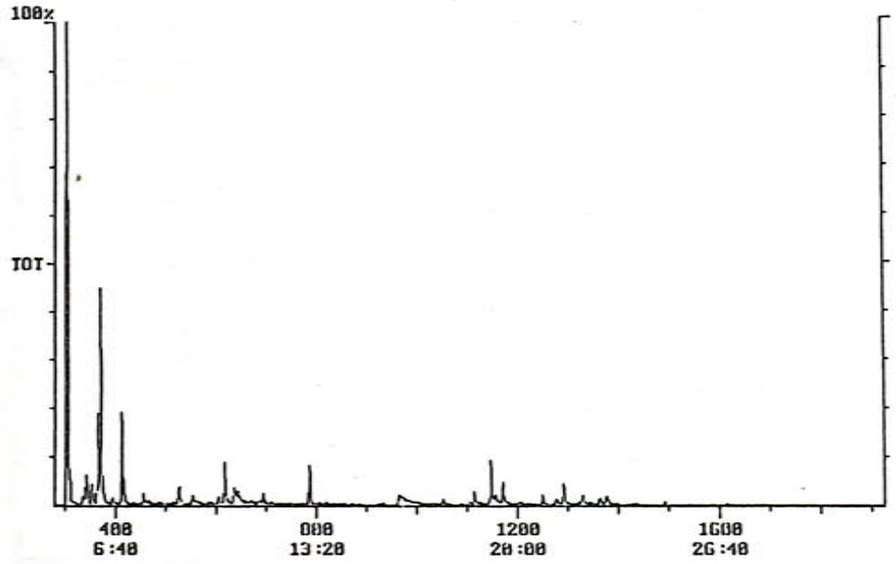
9


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

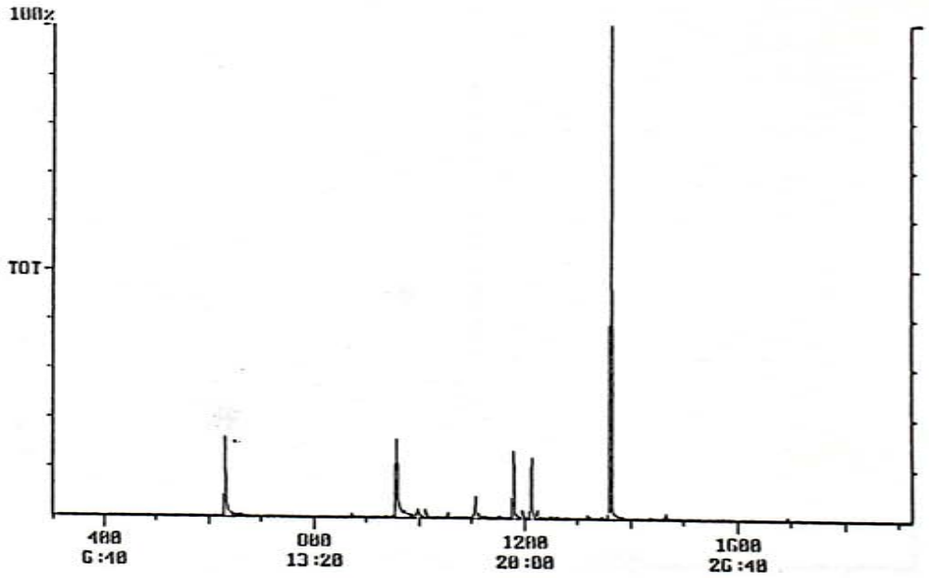
11


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

10

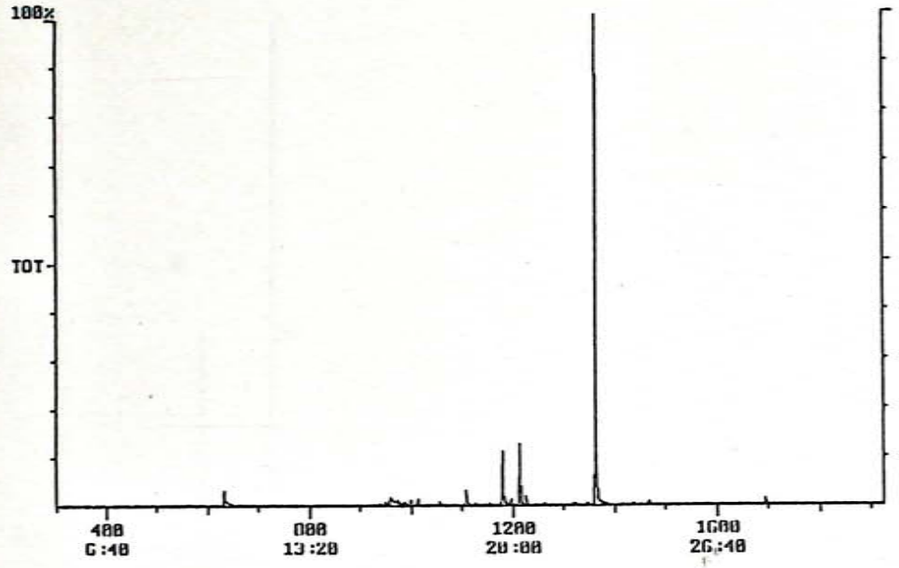

100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

12

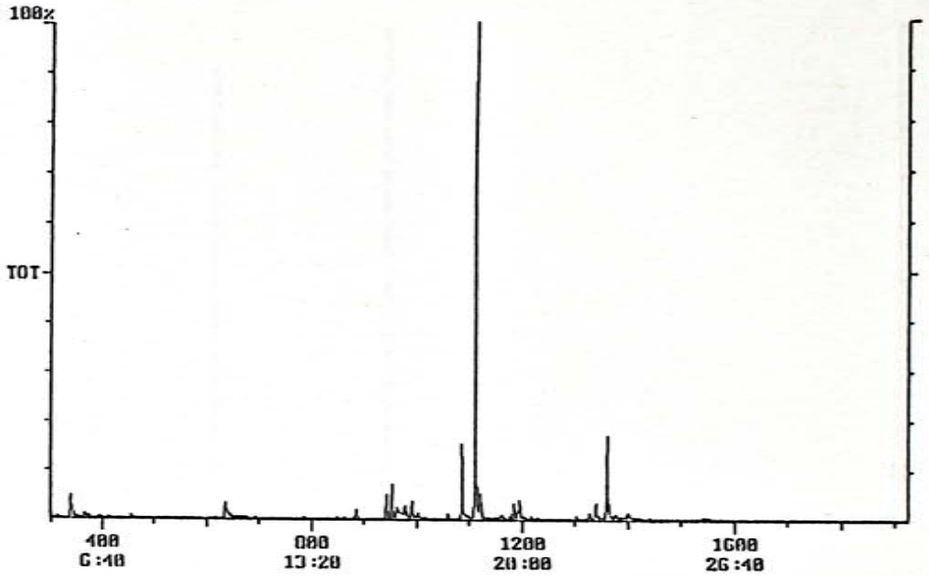

100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

A6

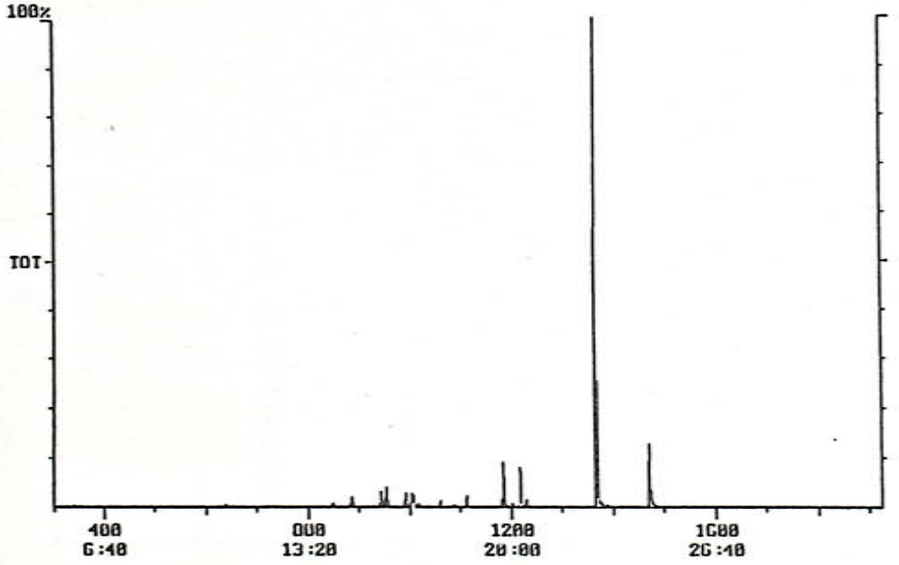
13



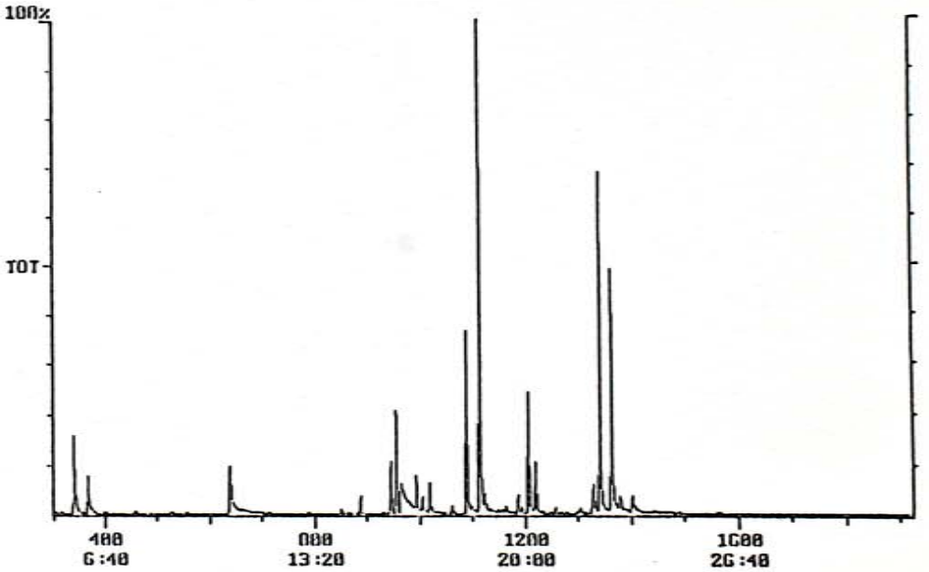
15



14

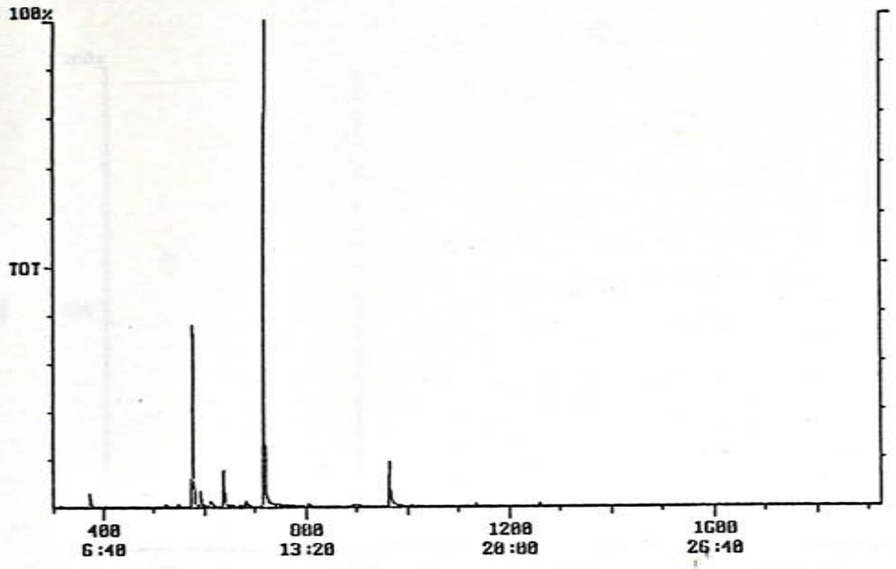


16

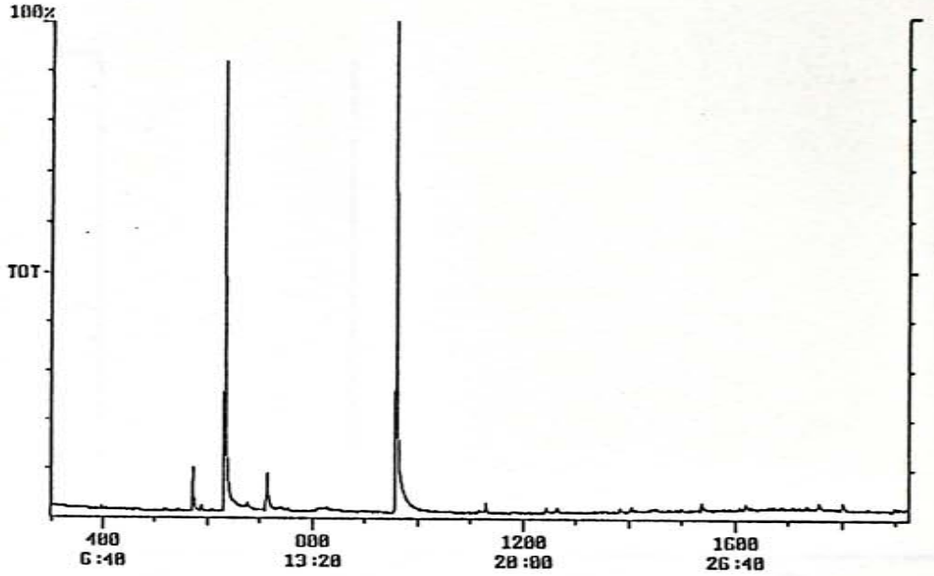


A11

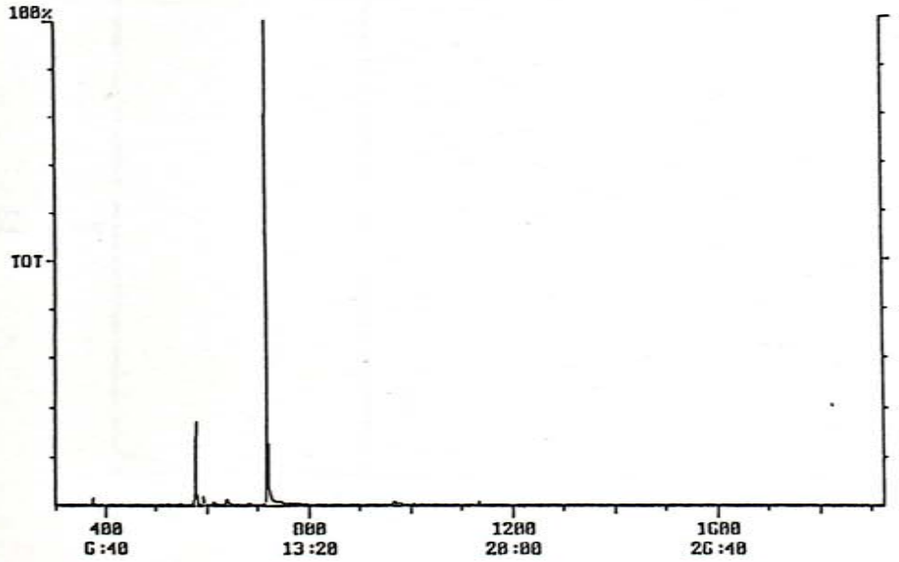
17



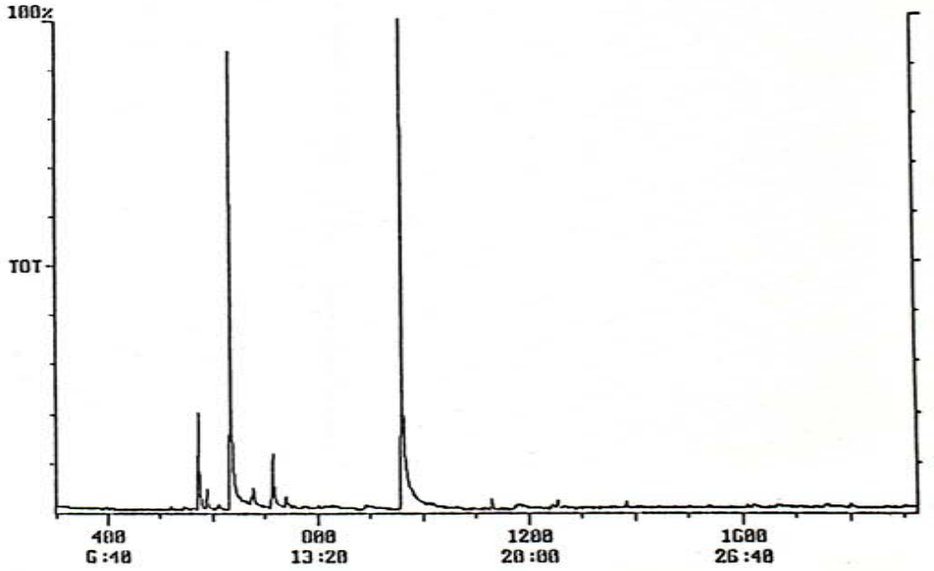
19



18

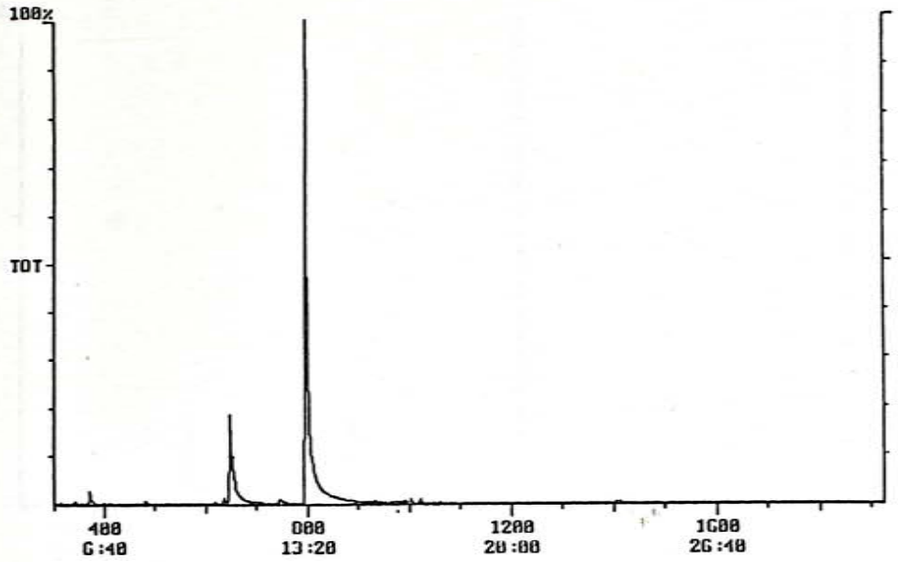


20

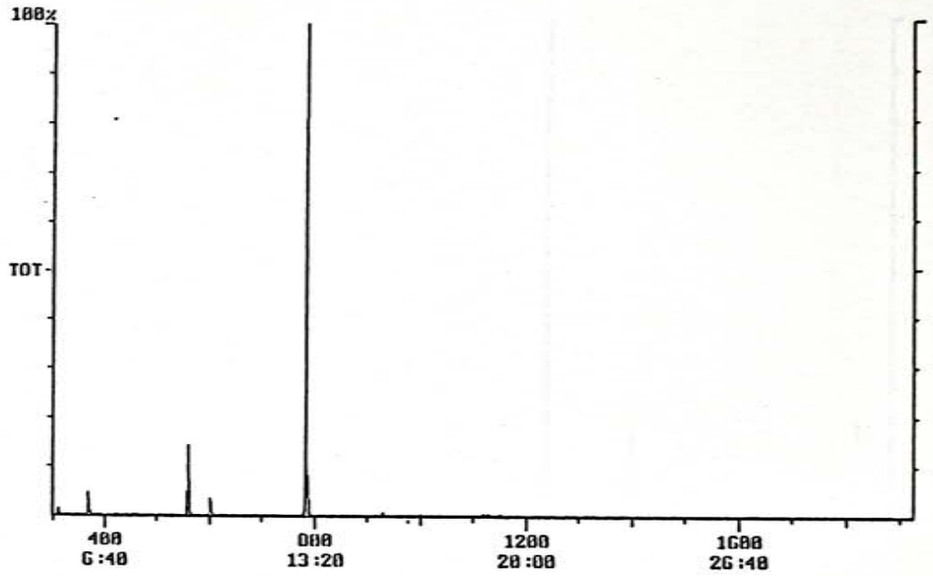


A12

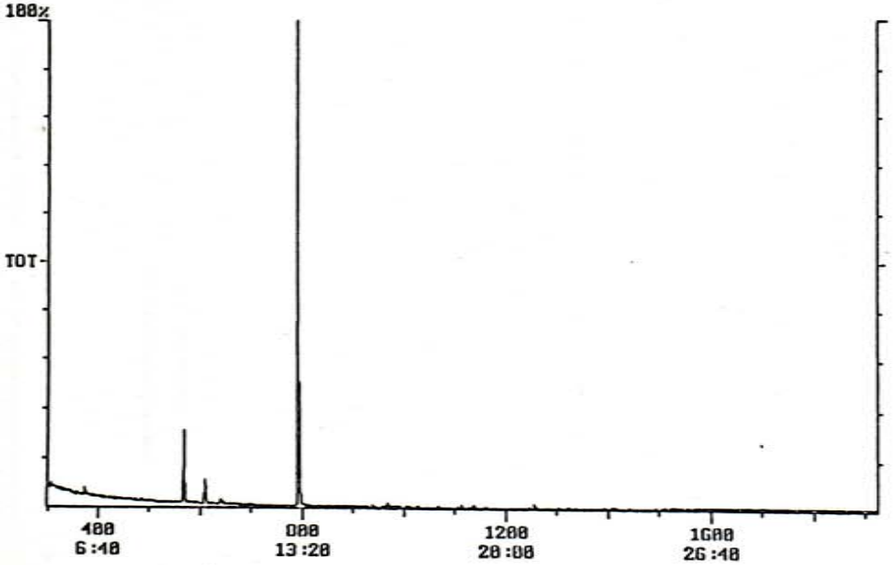
21



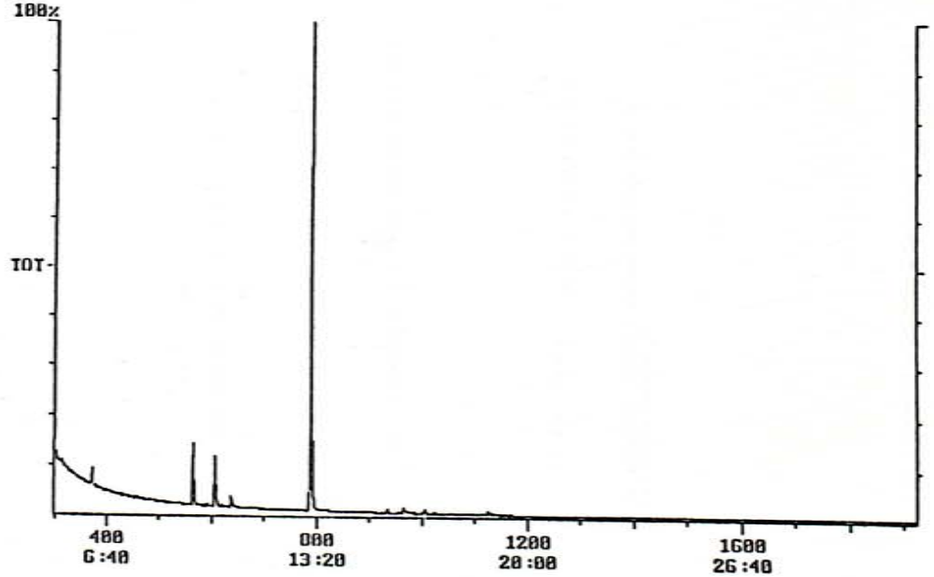
23



22

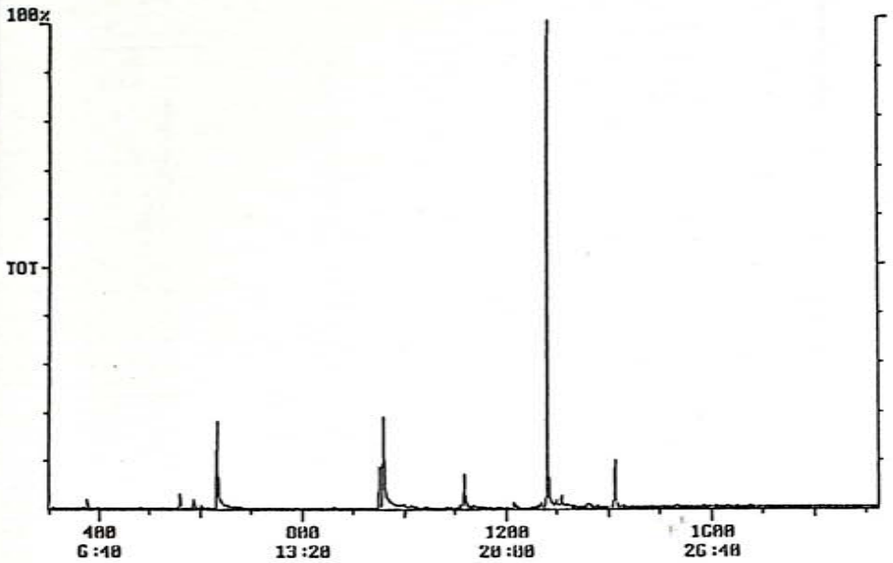


24

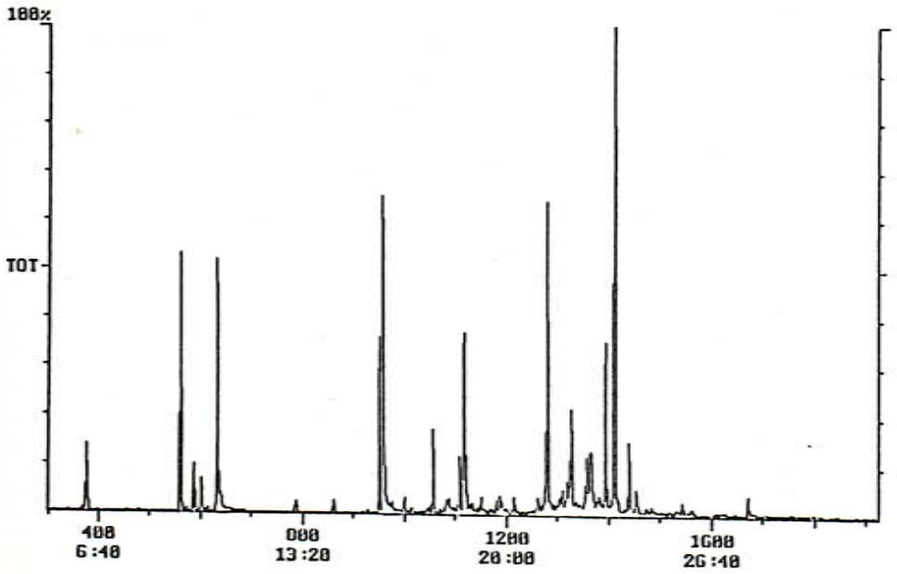


A13

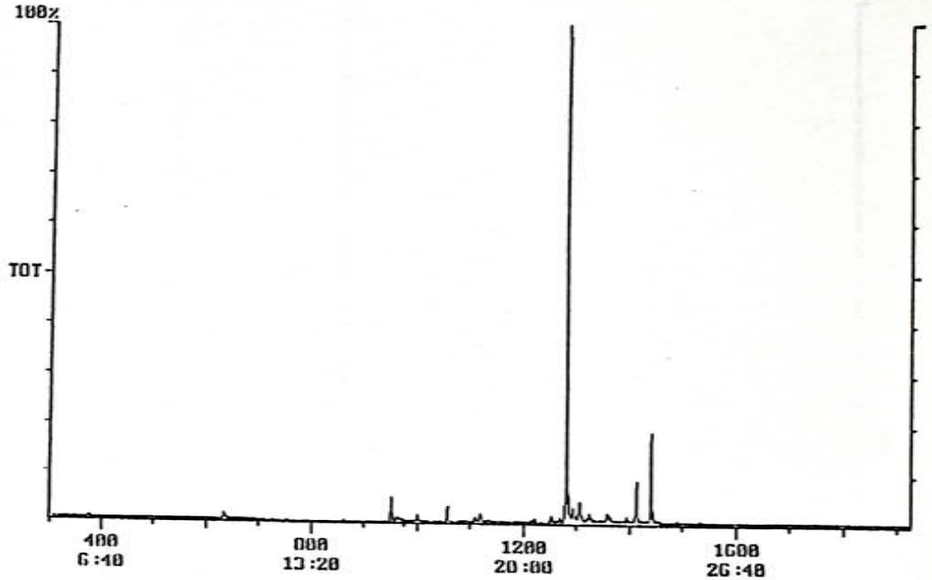
25



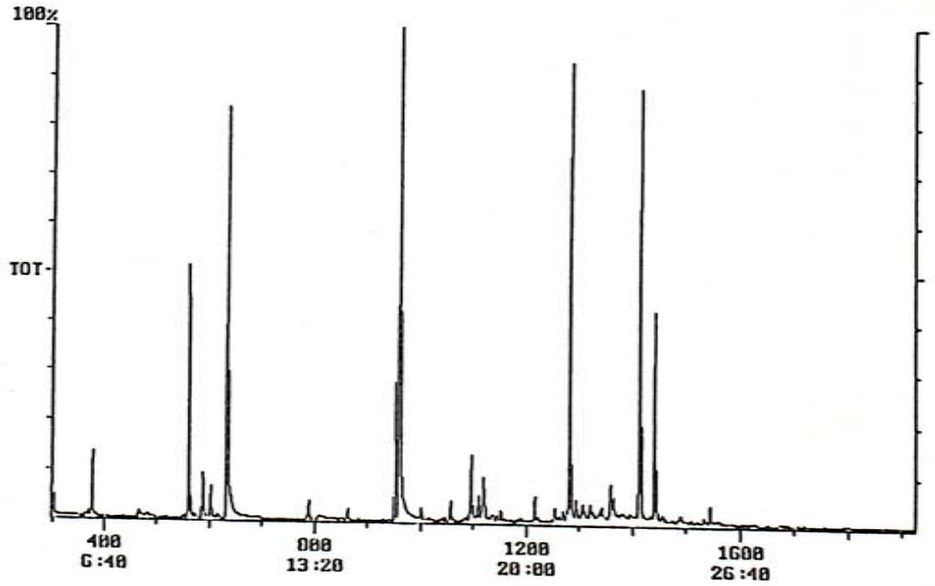
26



27

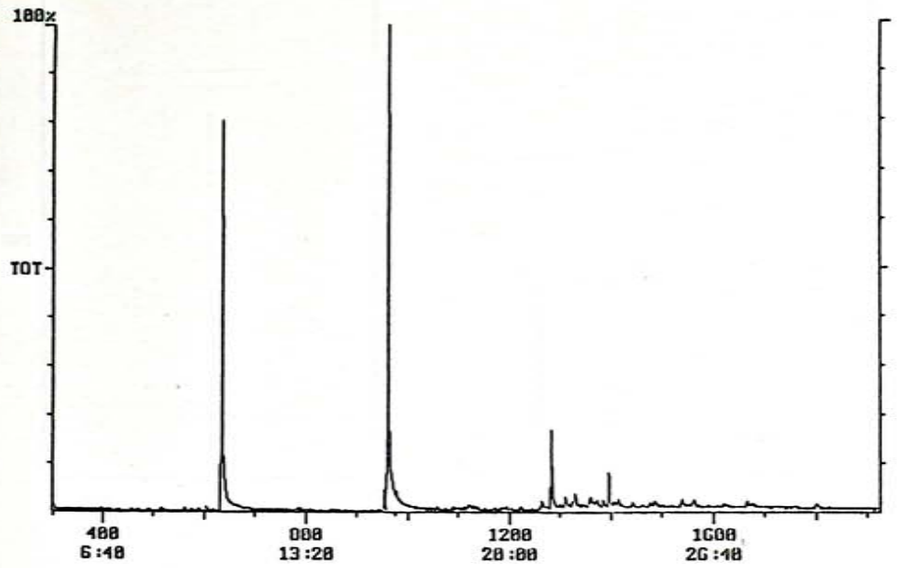


28

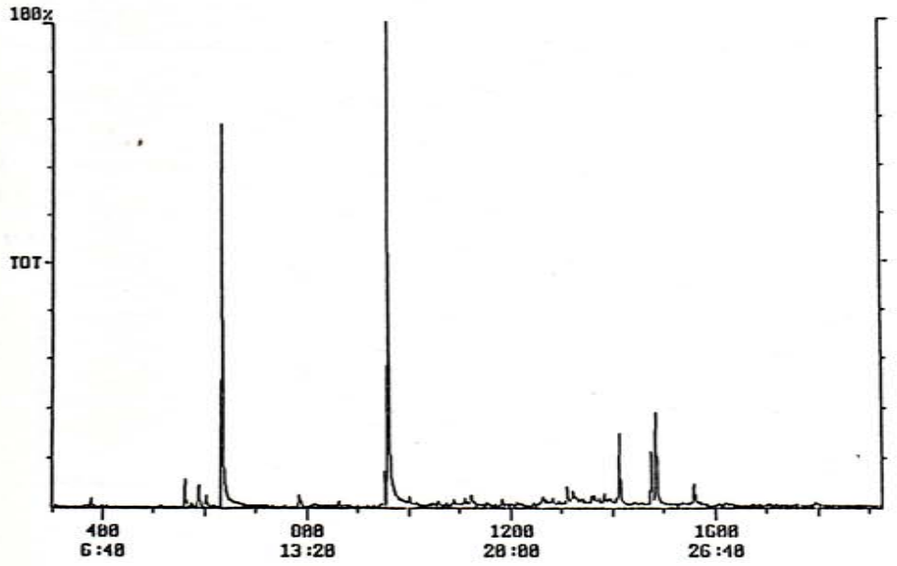


A14

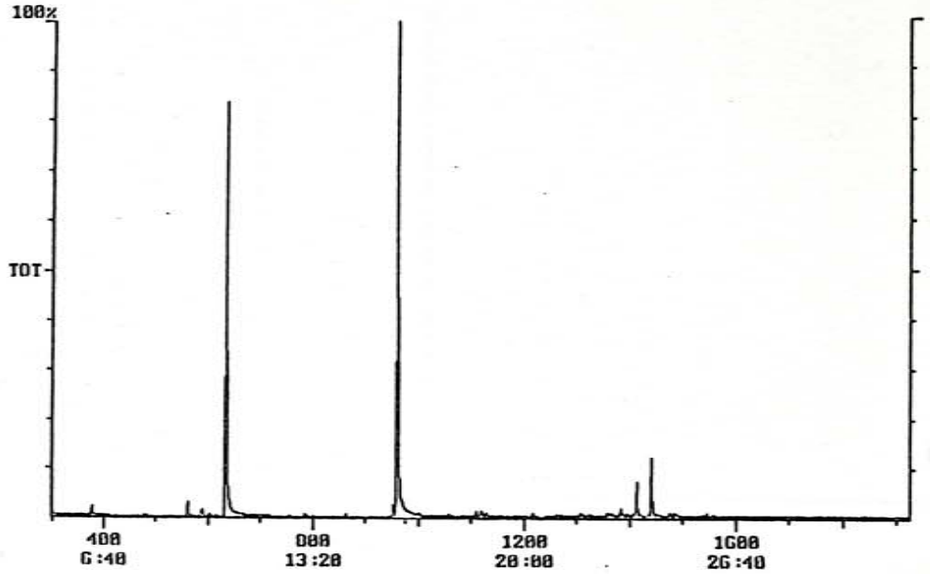
29



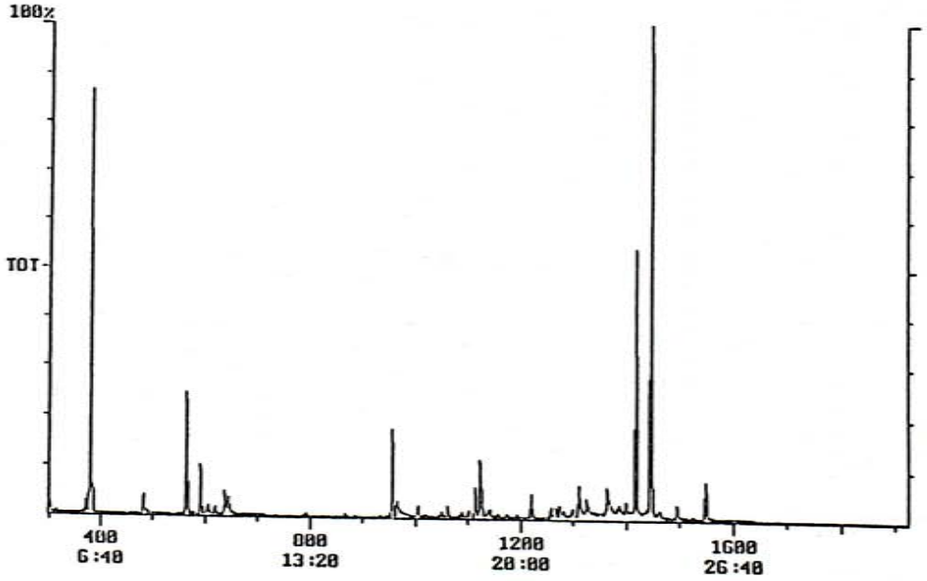
30



31

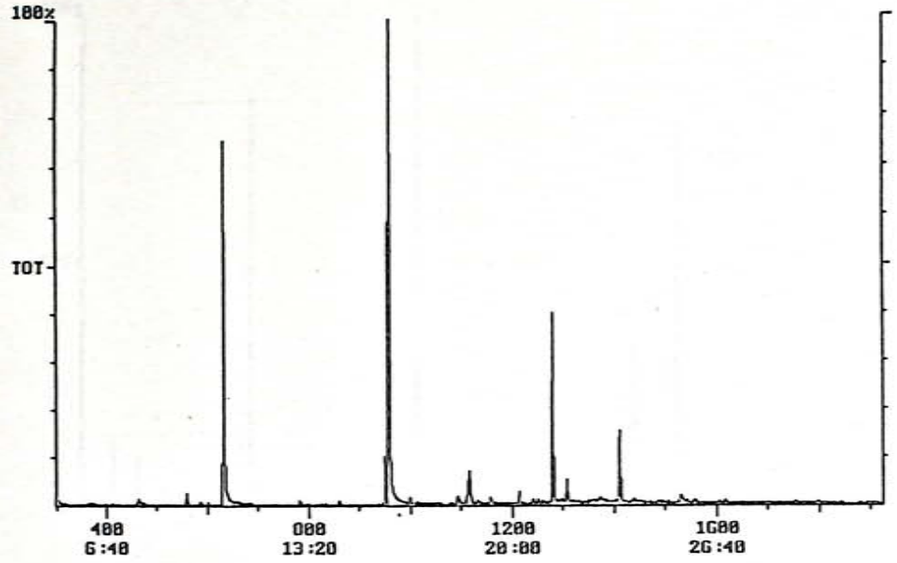


32

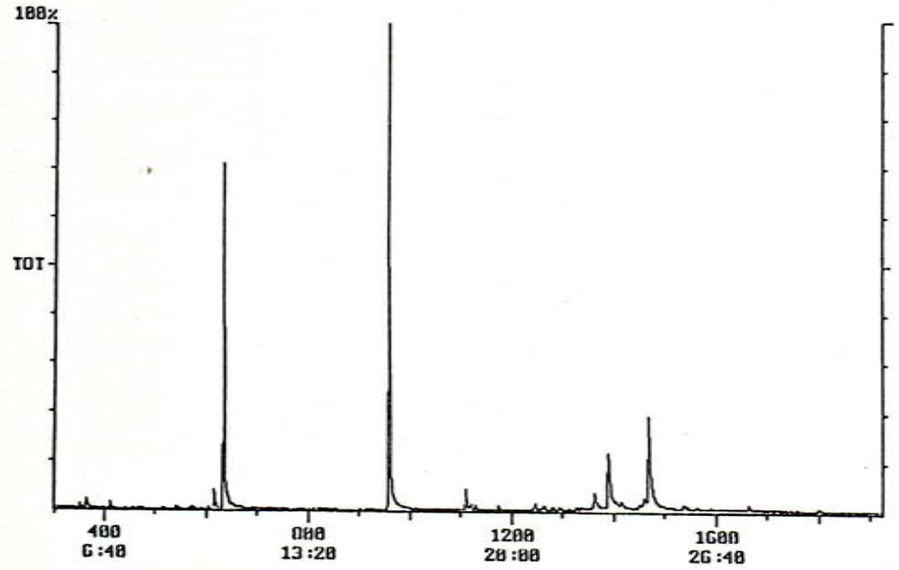


A15

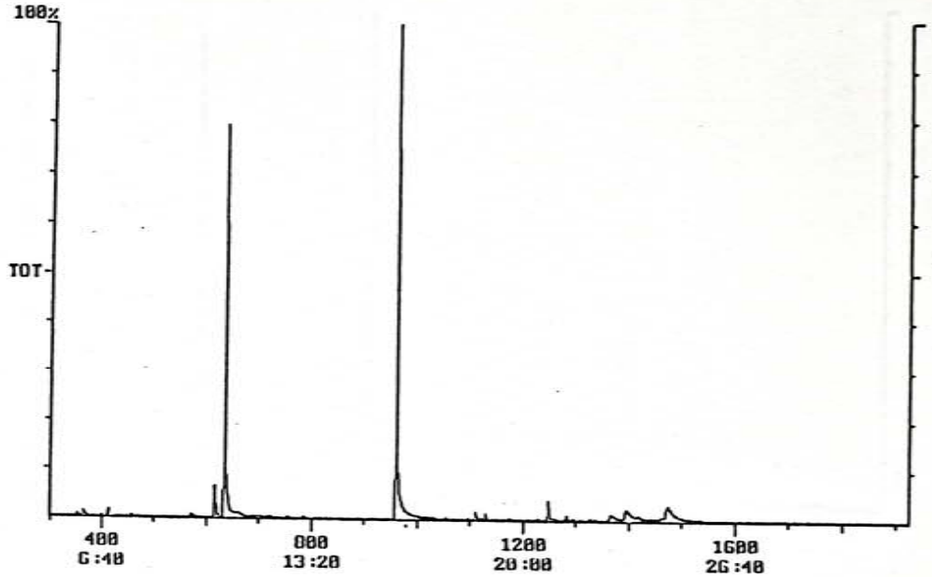
33


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

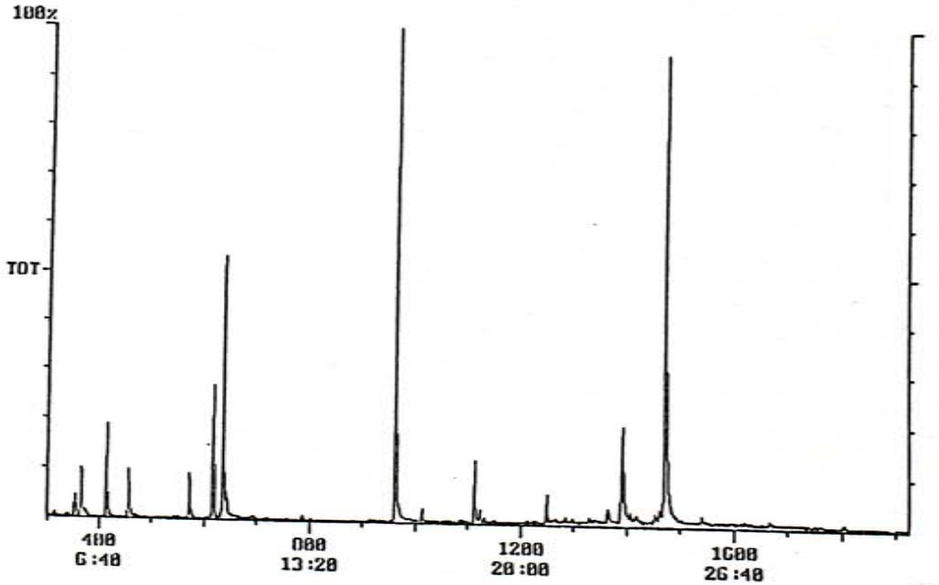
34


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

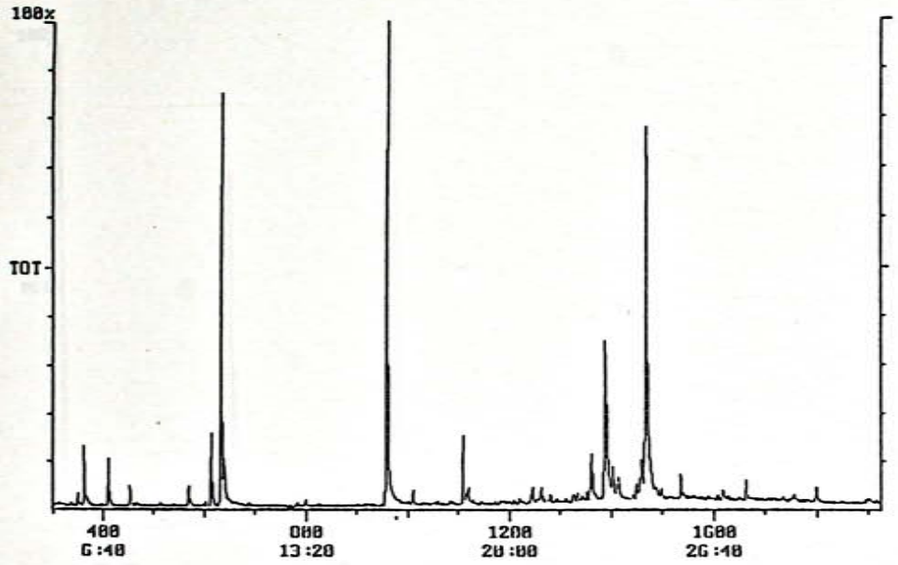
35


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

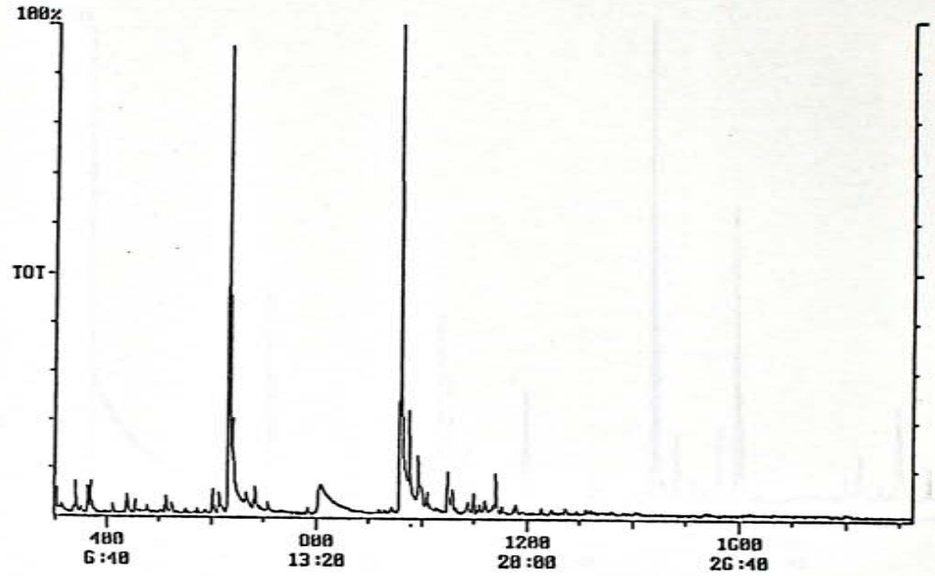
36


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

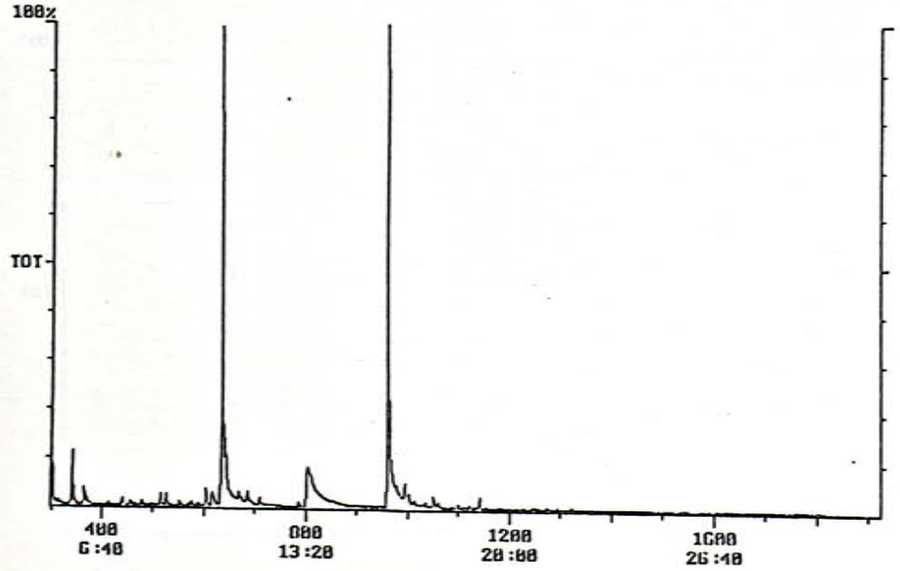
37


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

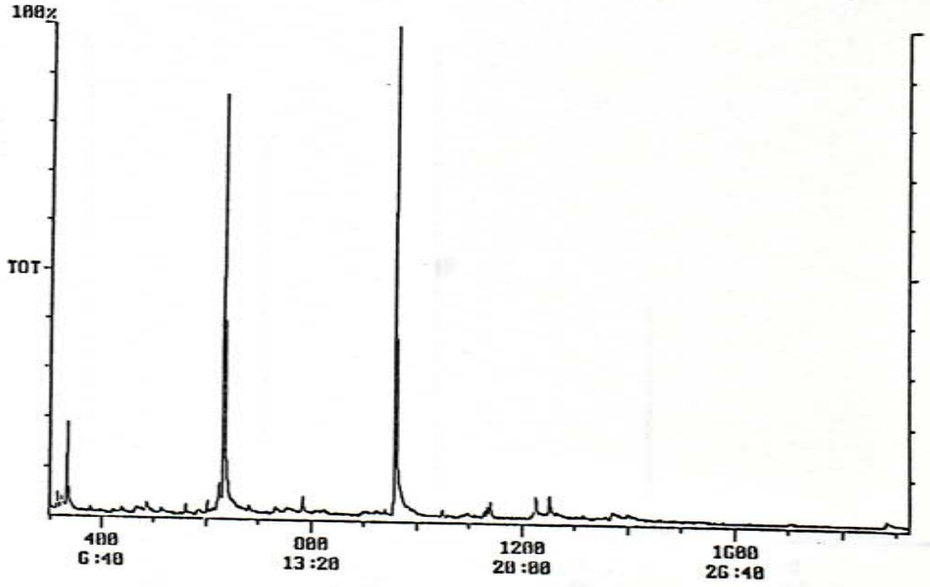
39


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

38

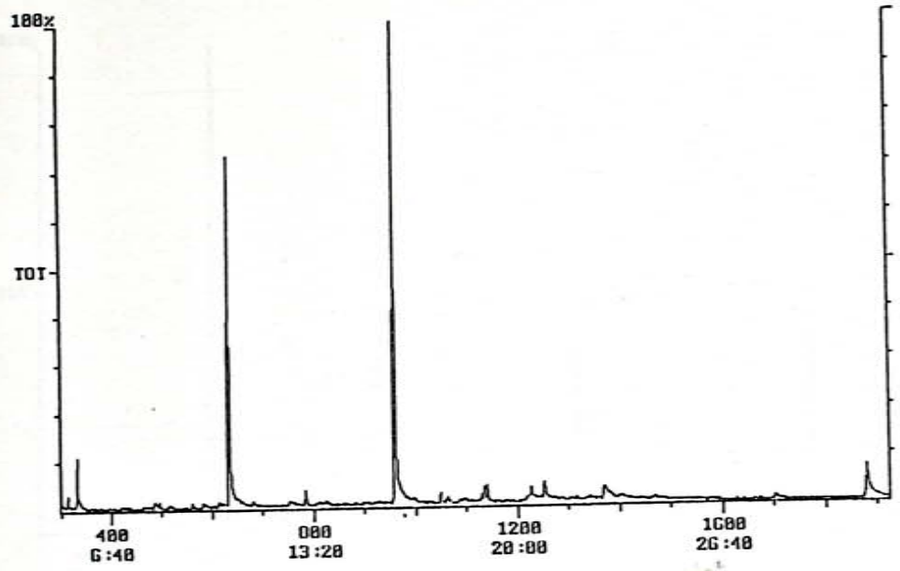

100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

40

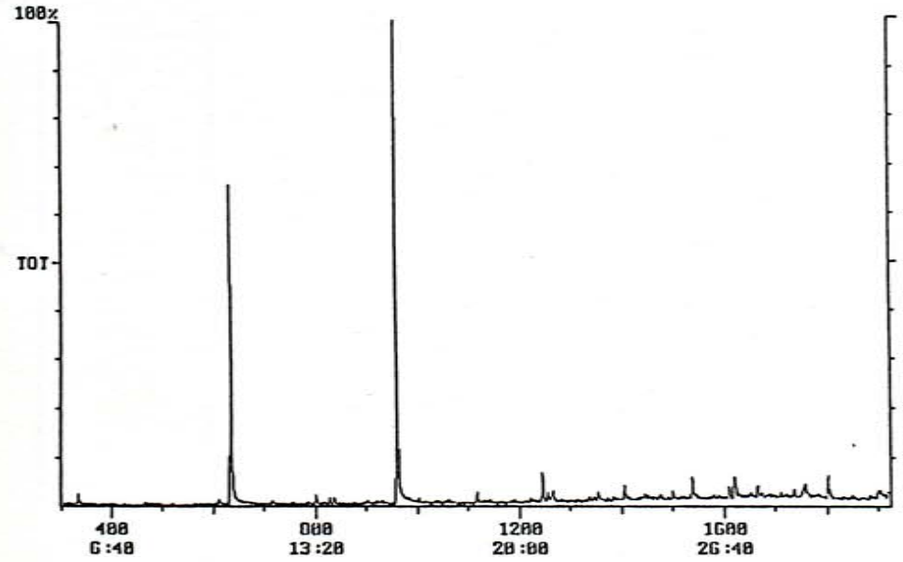

100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

A-7

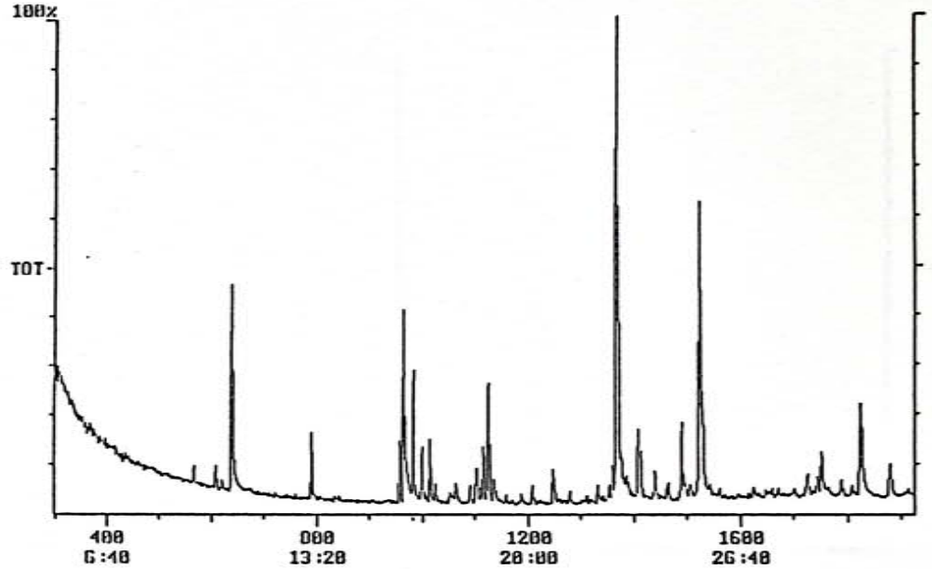
41



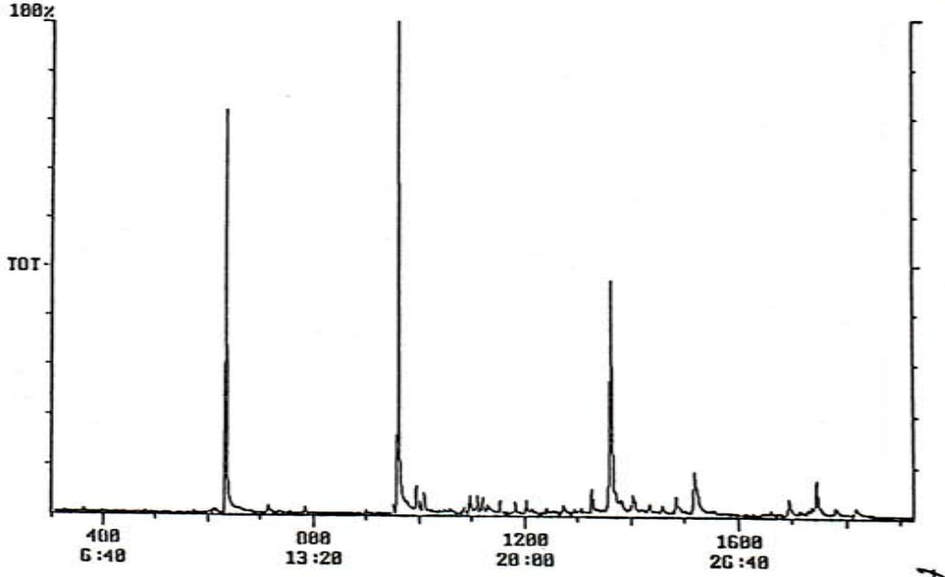
42



43

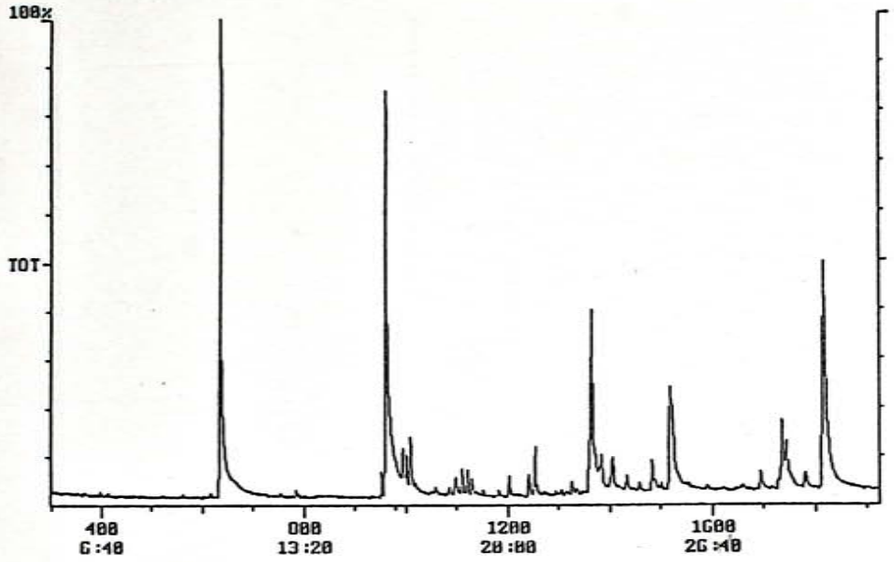


44

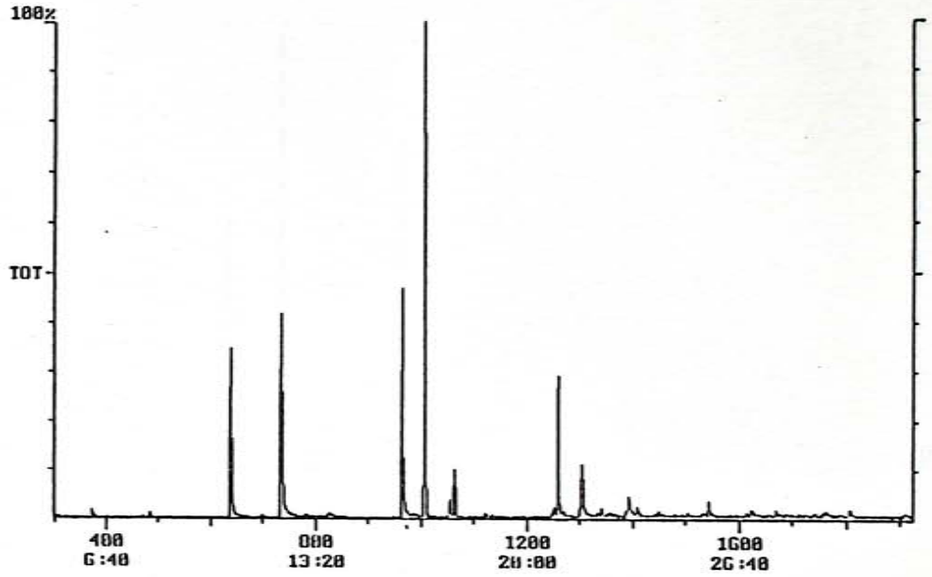


A18

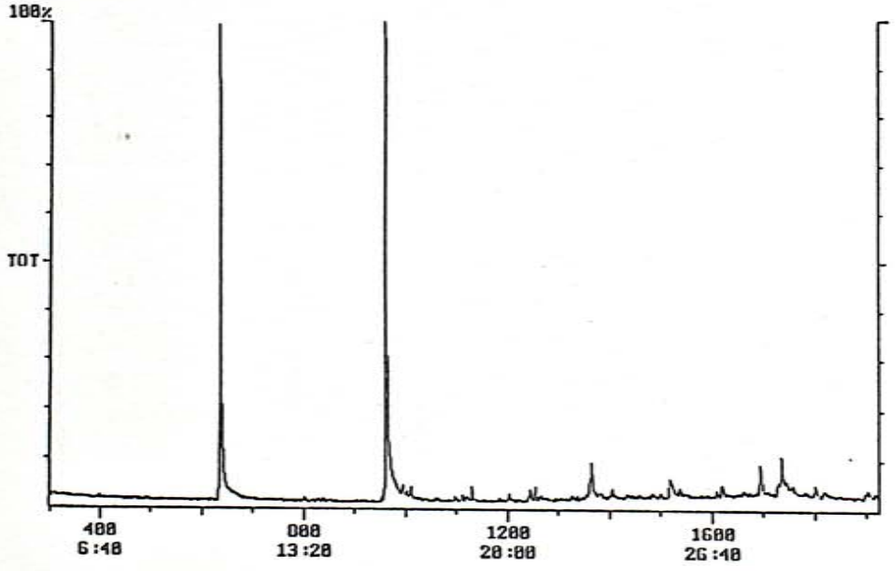
45


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

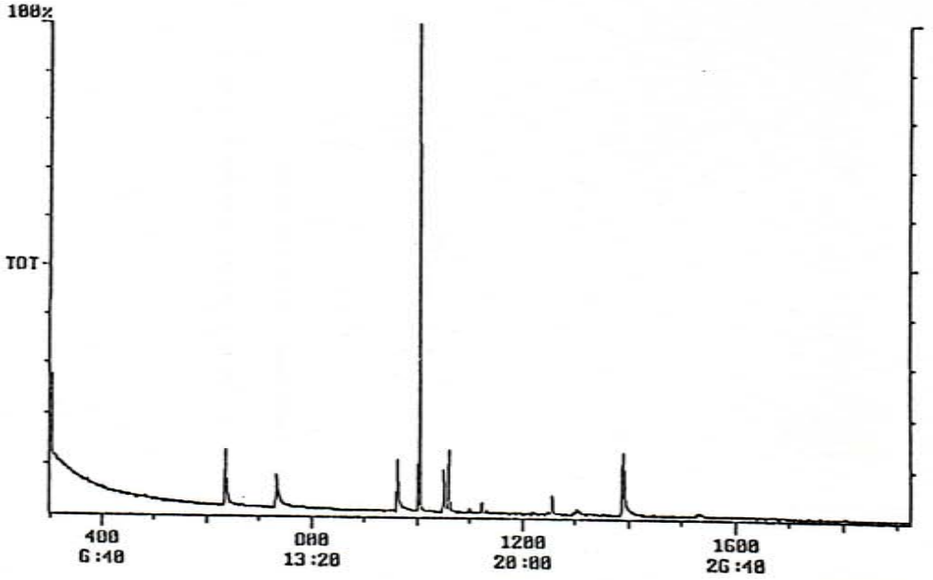
47


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

46

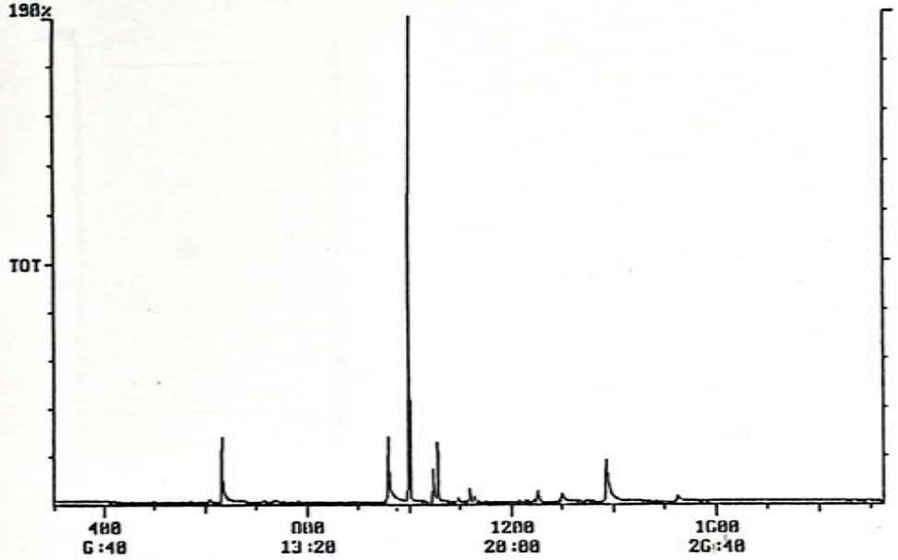

100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

48

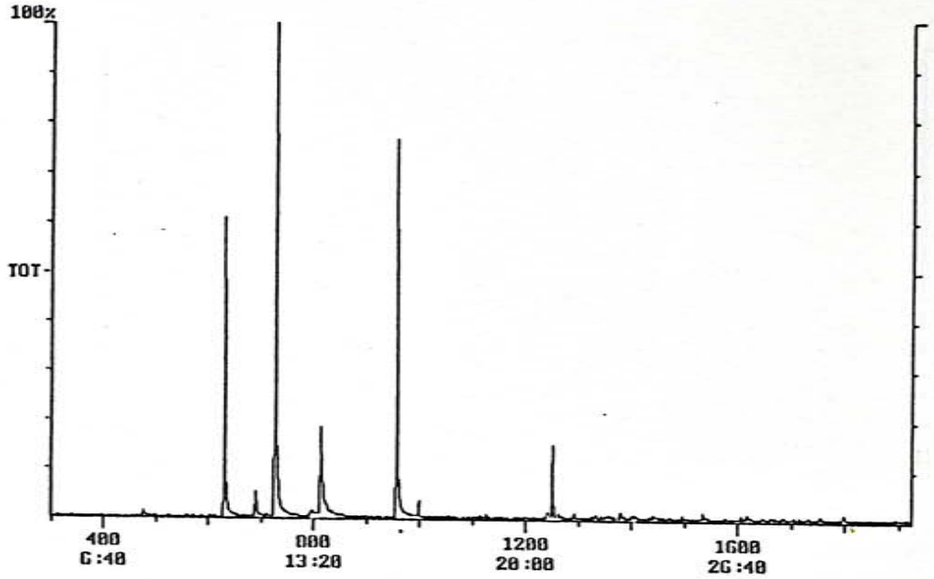

100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

49

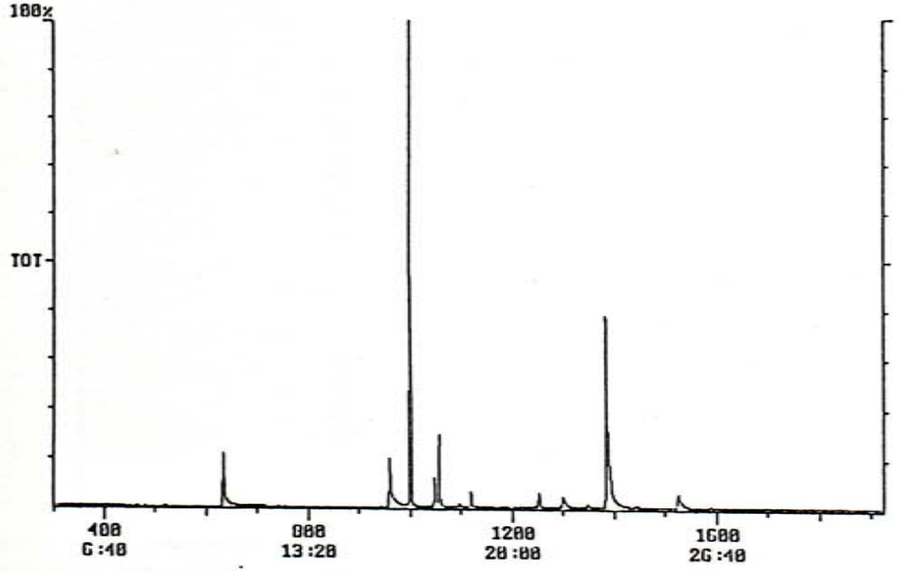
49



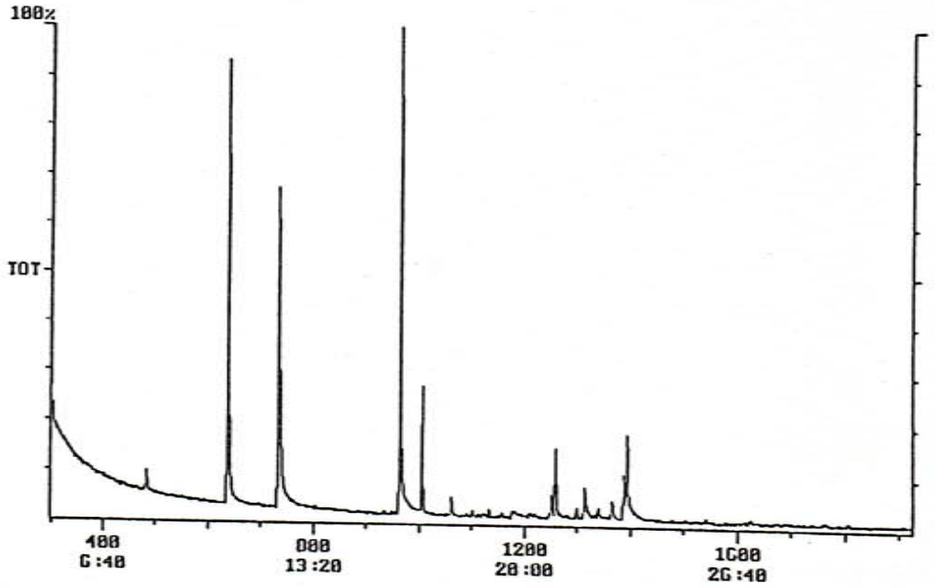
51



50

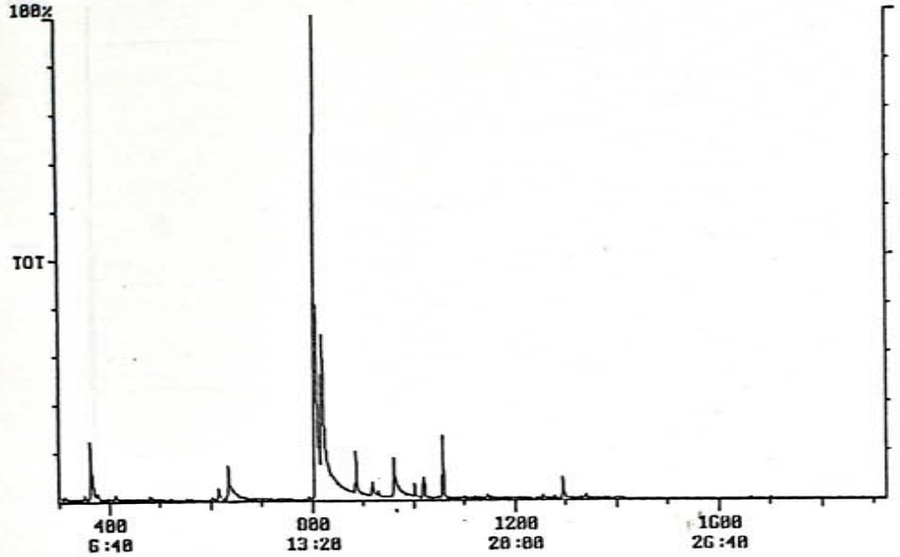


52

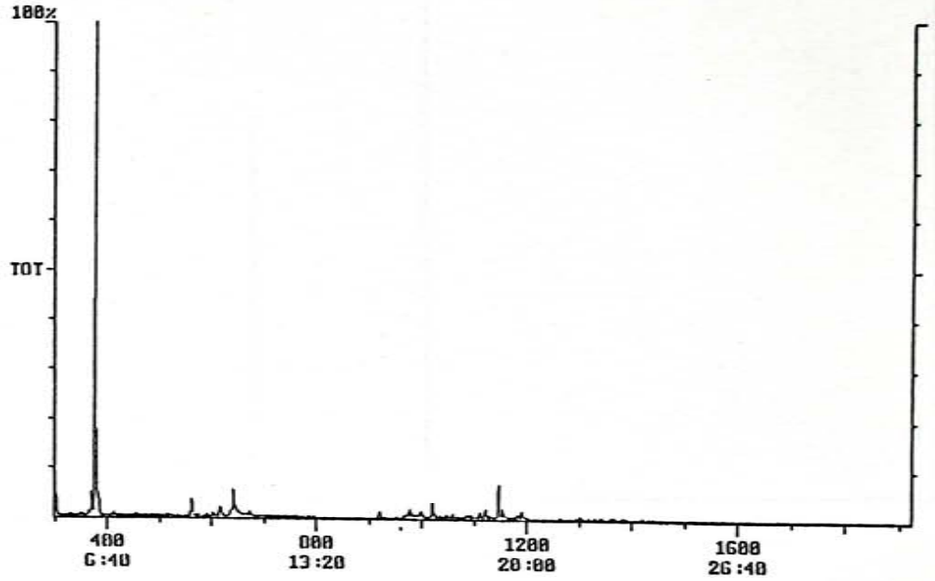


A26

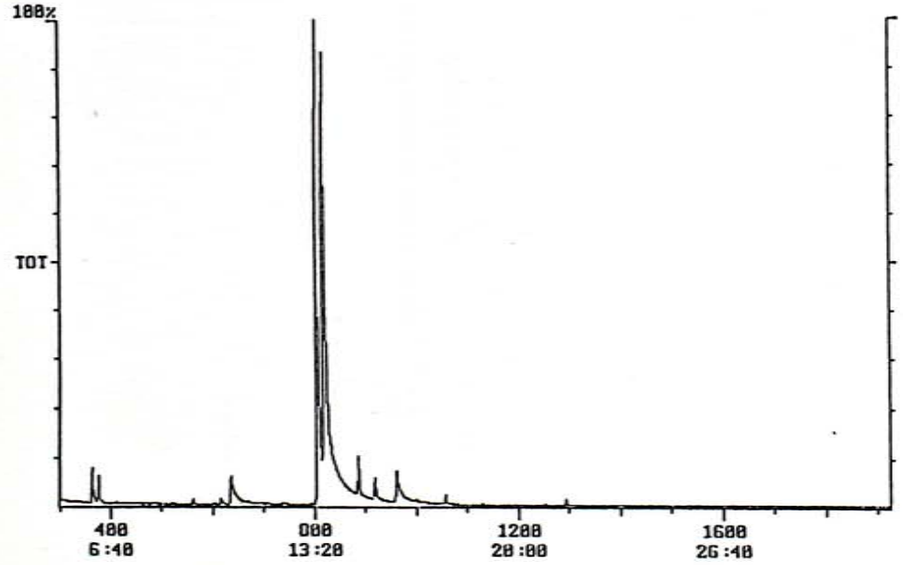
53



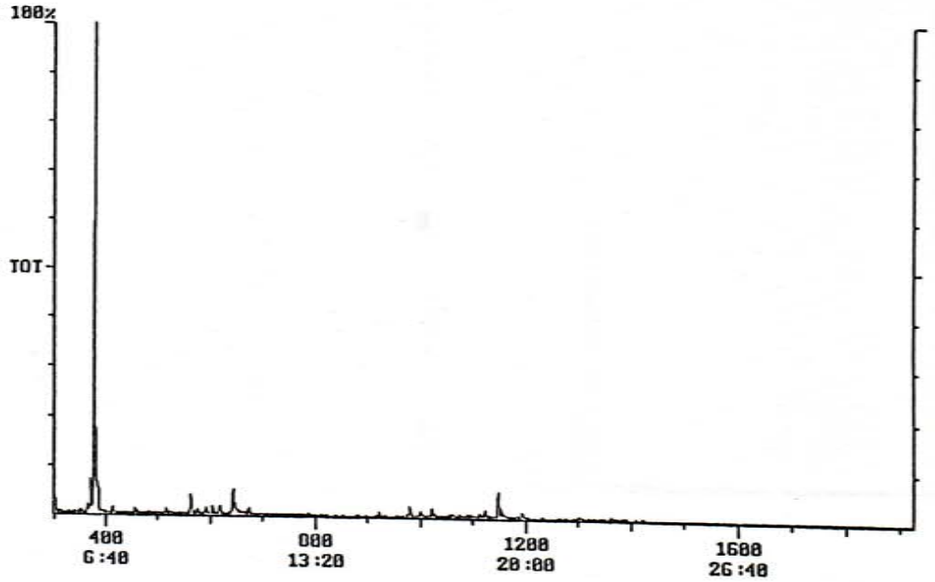
55



54

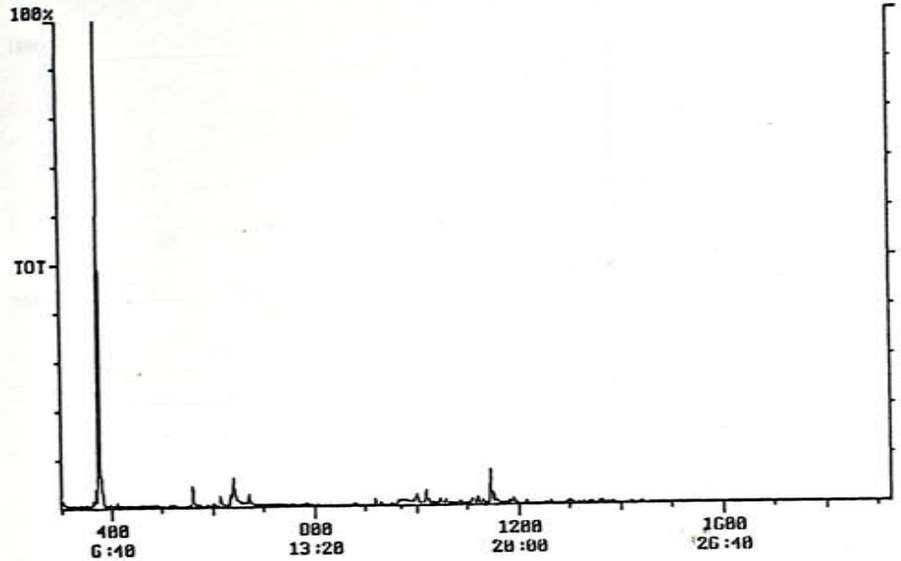


56

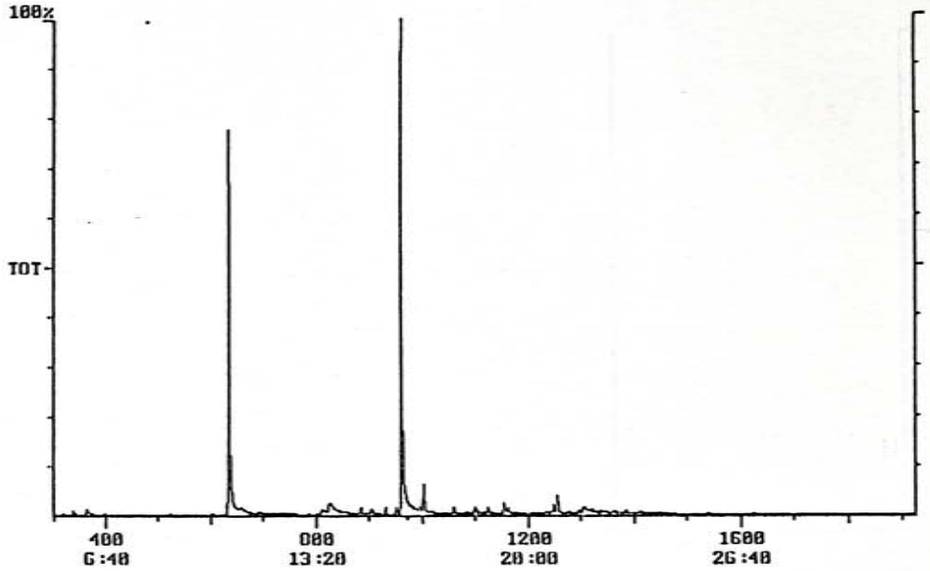


A21

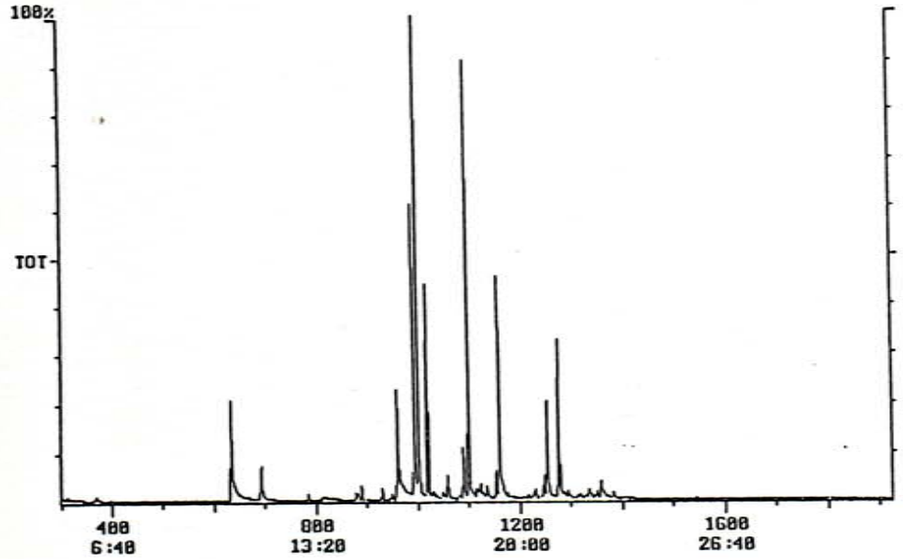
57



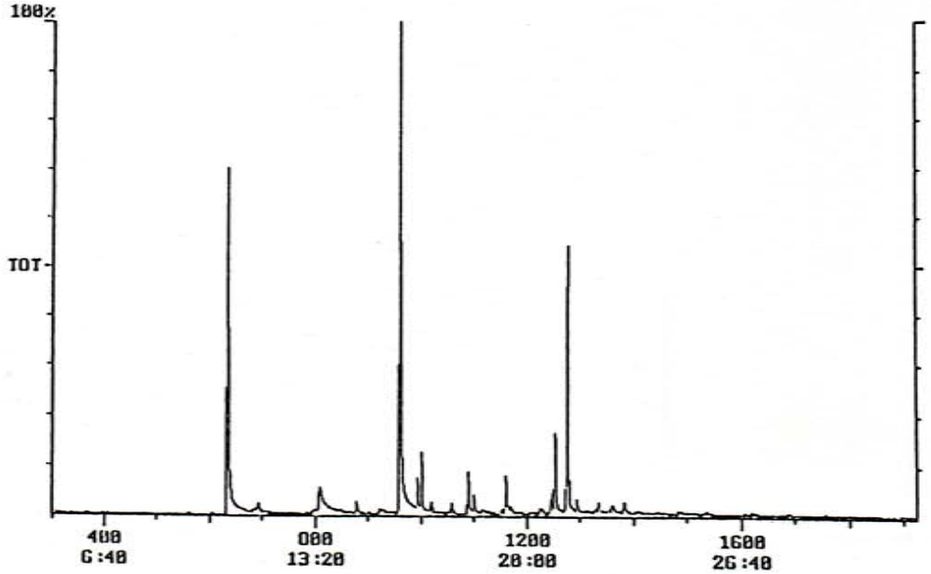
59



58

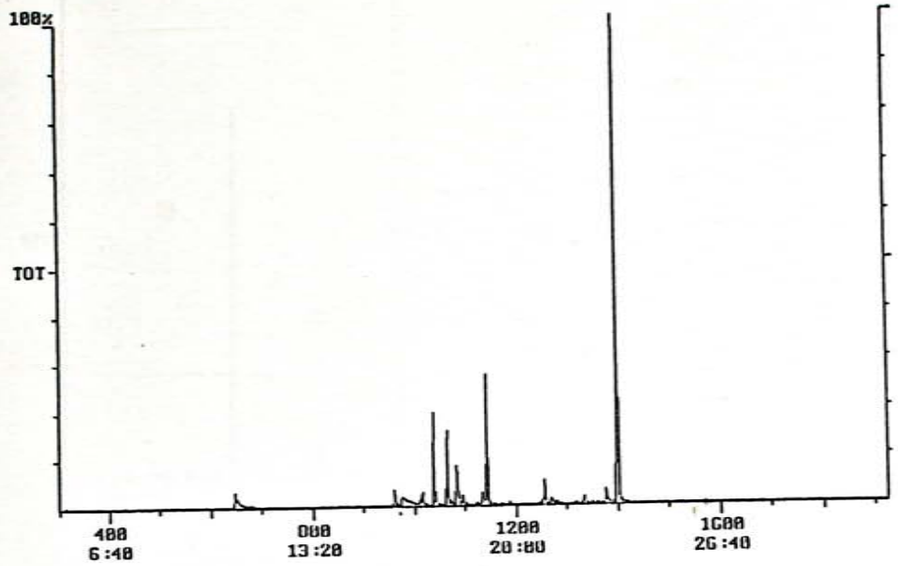


60

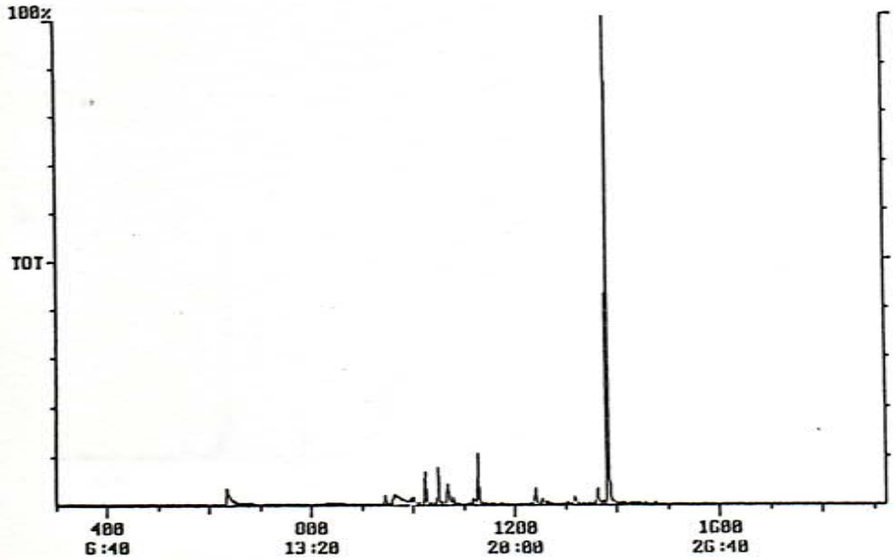


A22

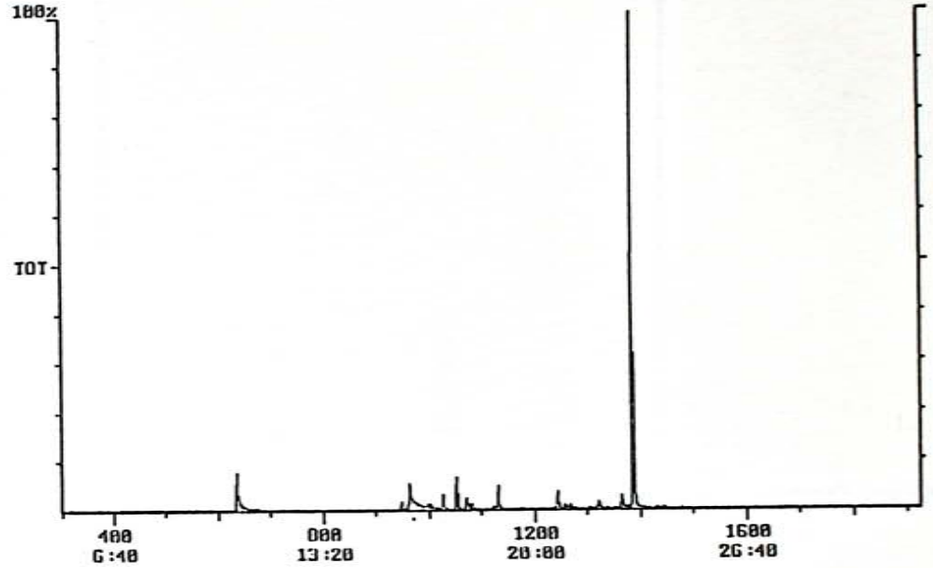
61


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

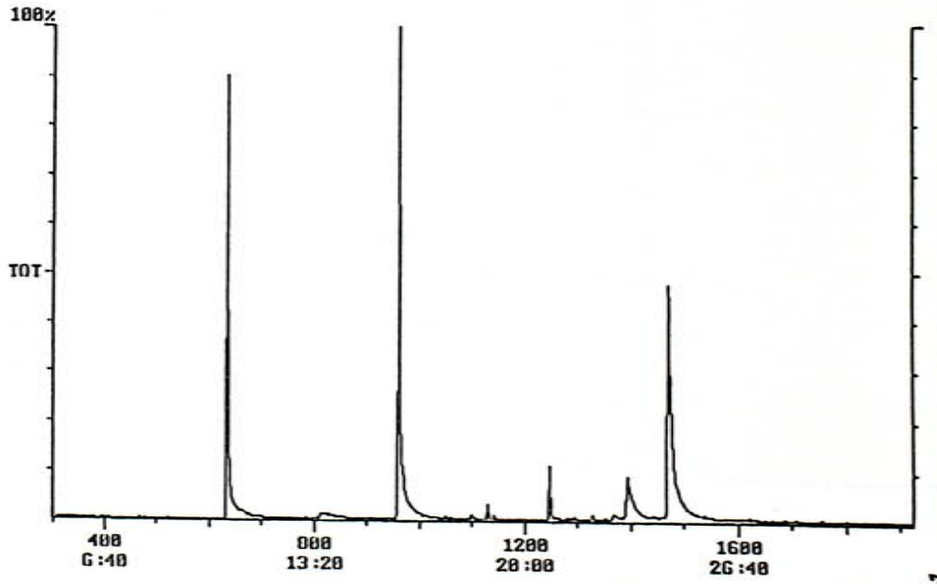
62


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

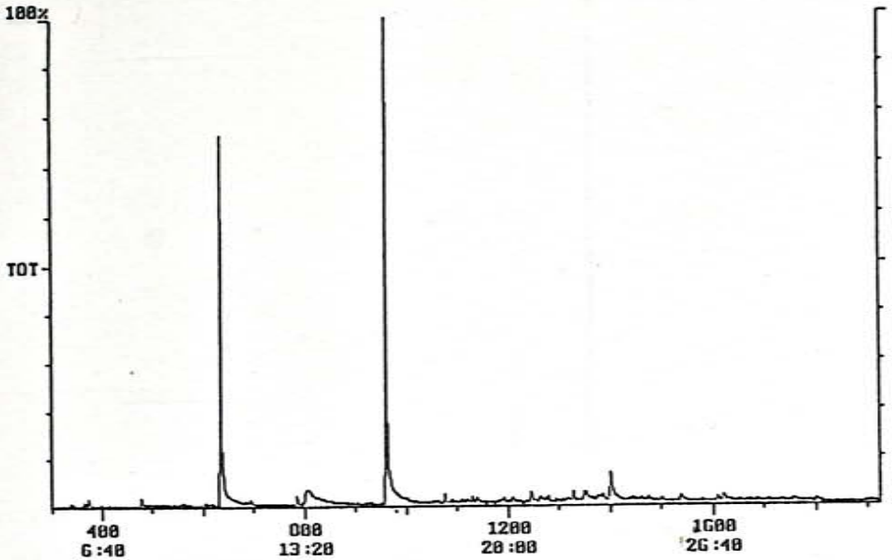
63


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

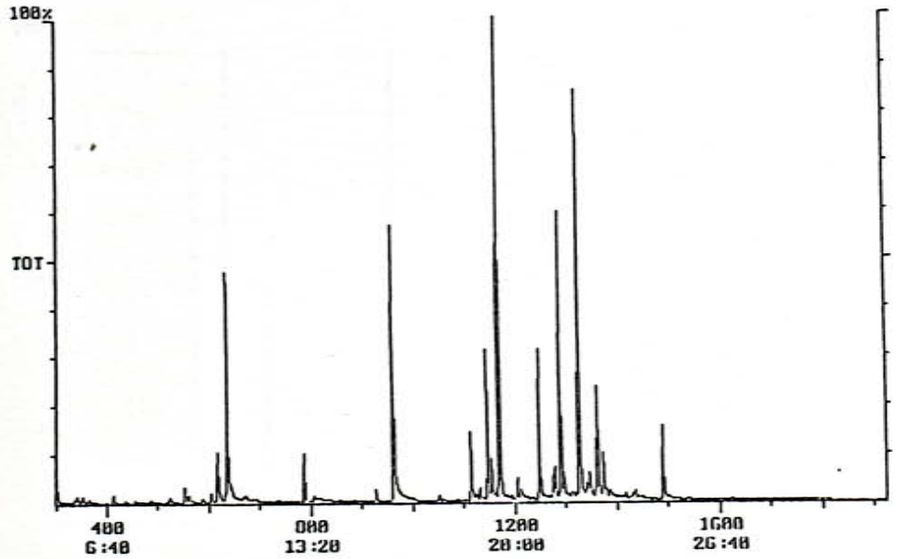
64


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

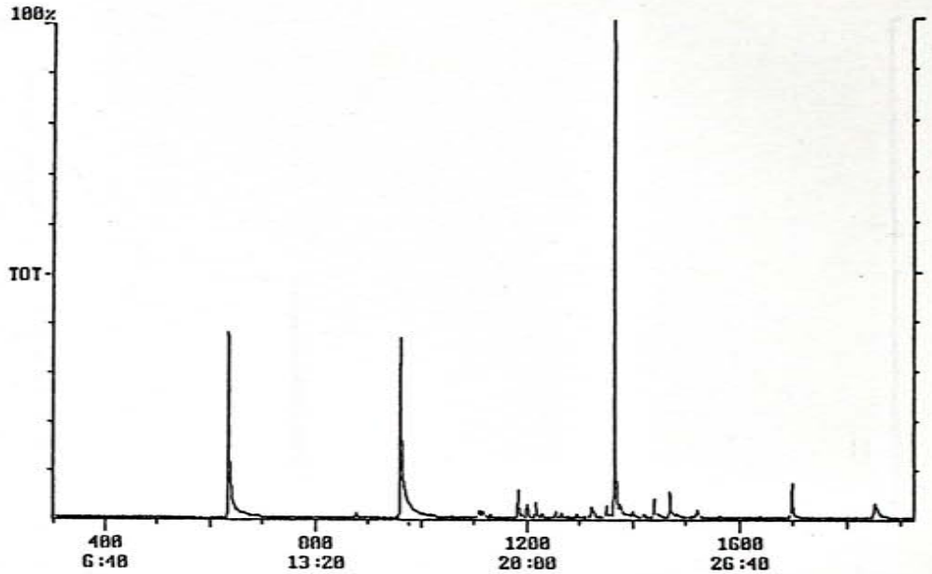
65


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

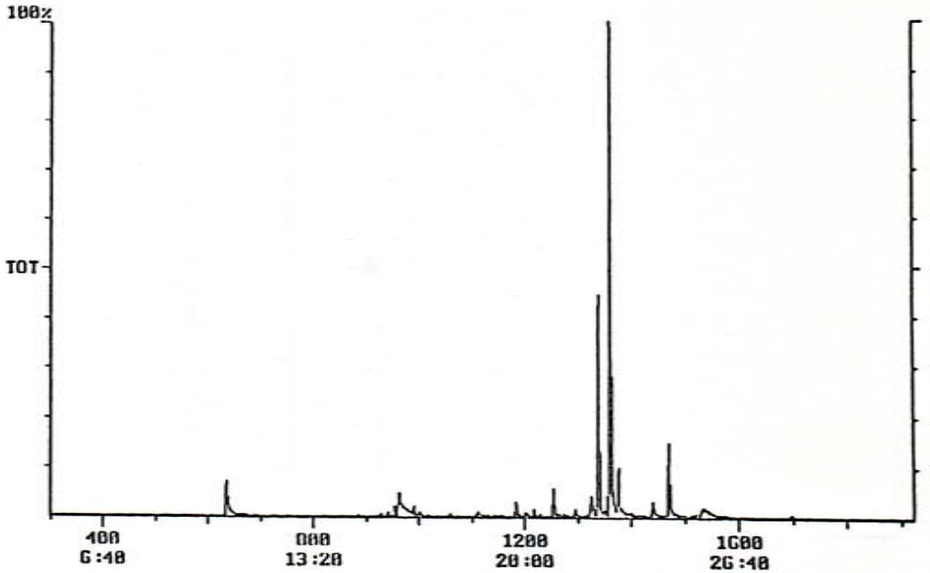
66


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

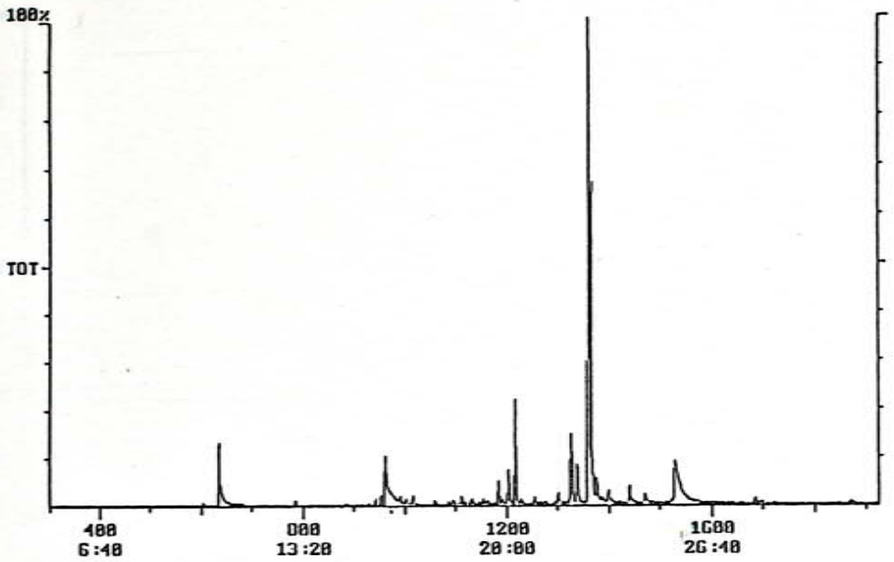
67


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

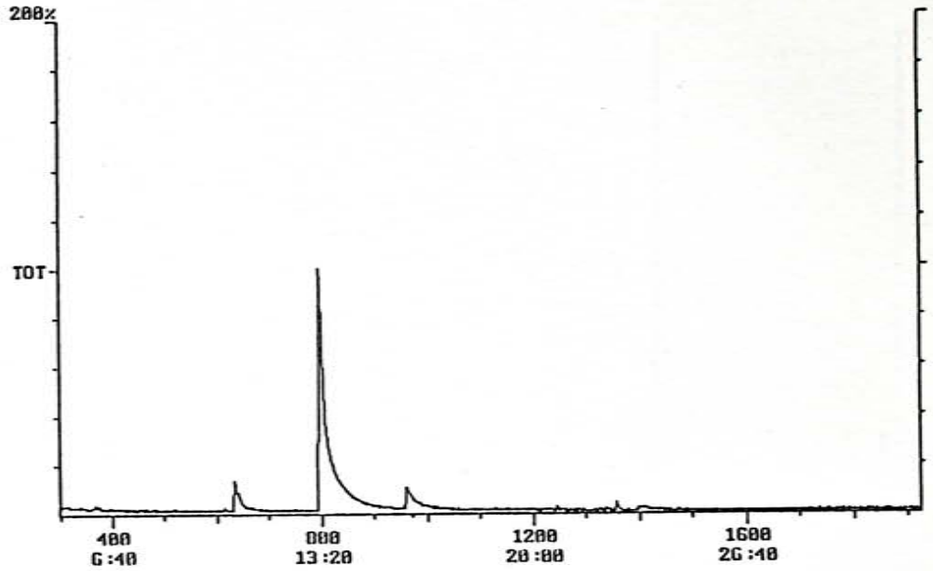
68


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

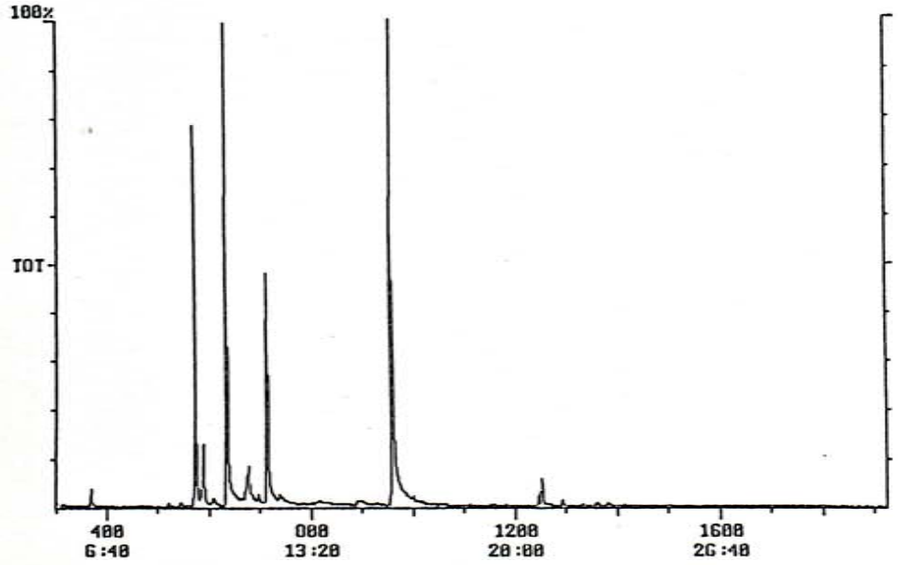
69



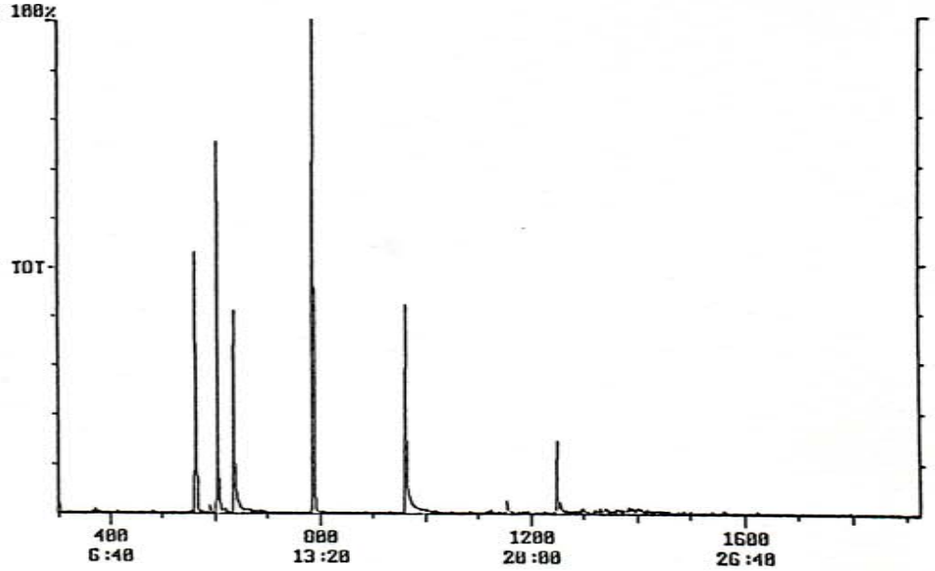
71



70

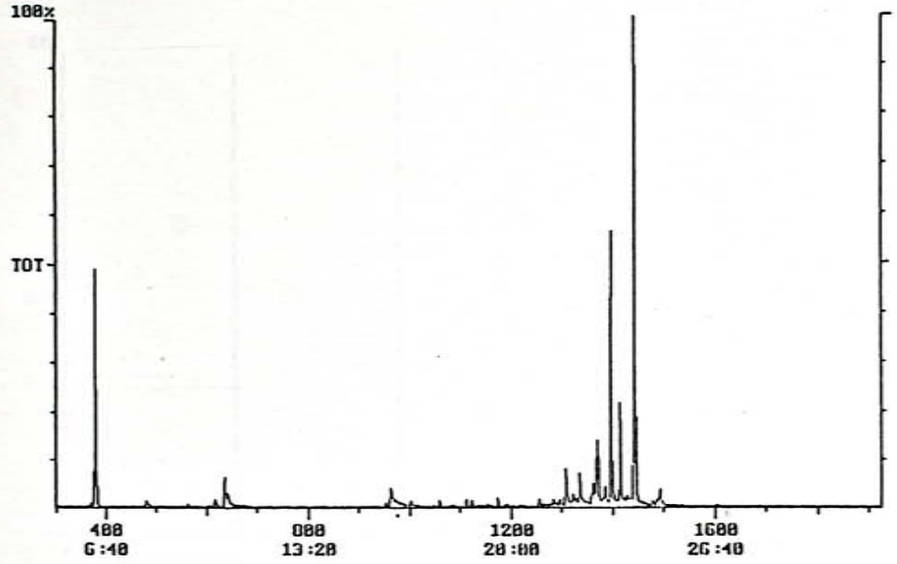


72

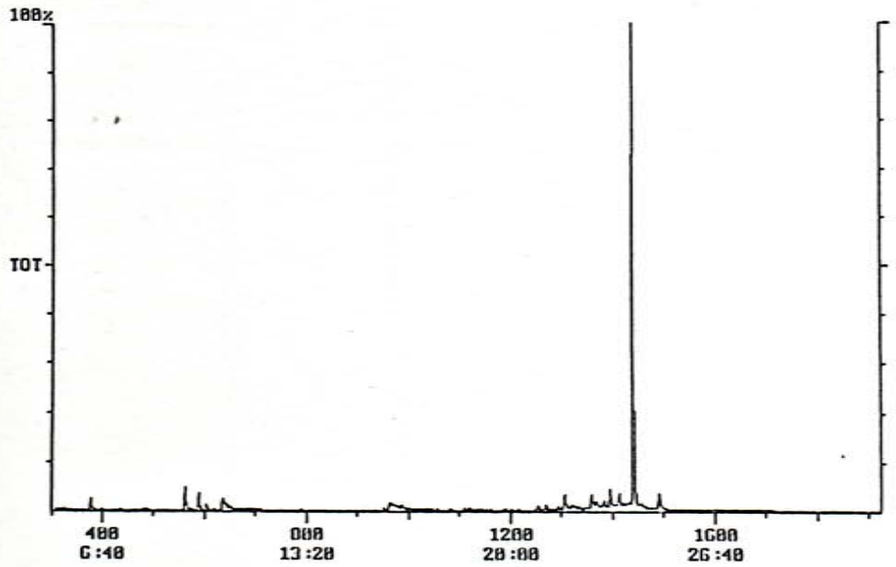


A25

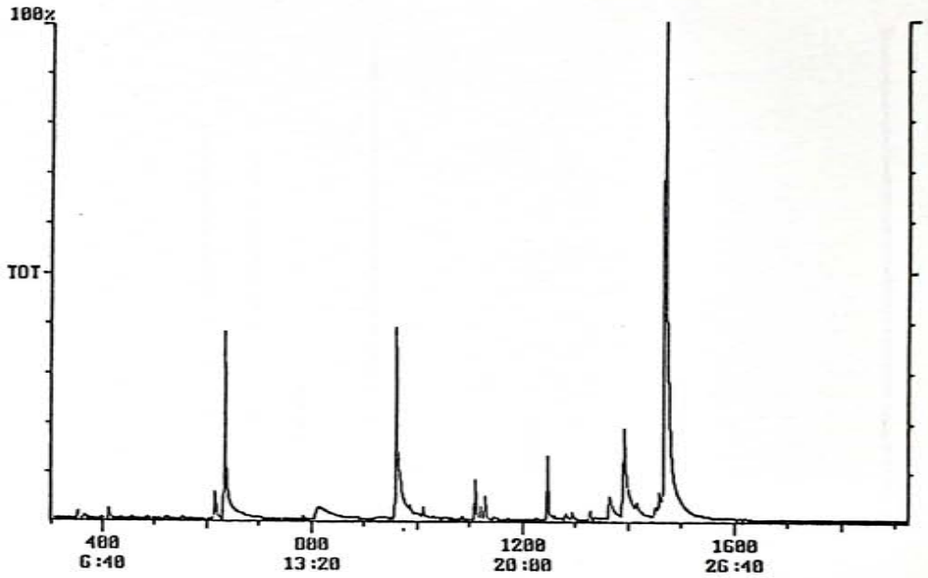
73


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

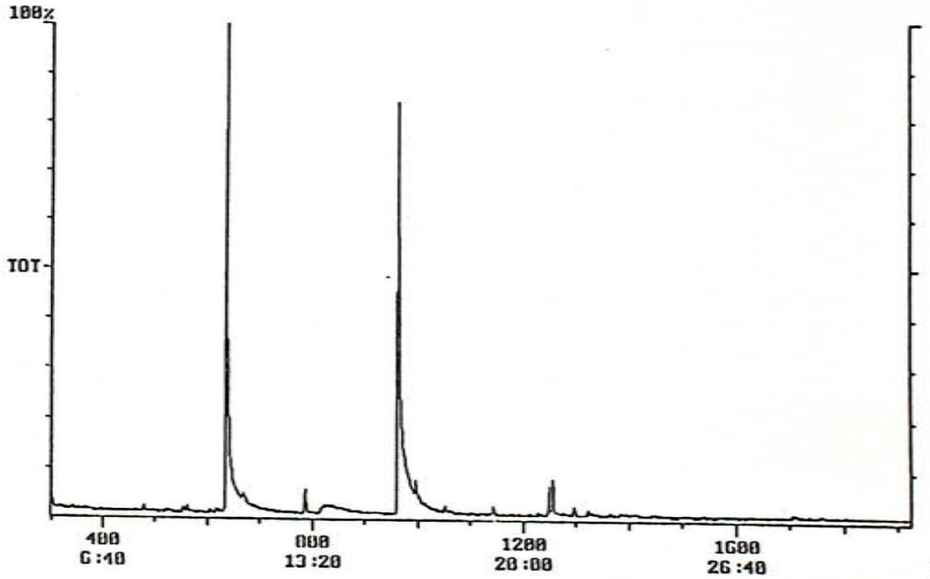
74


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

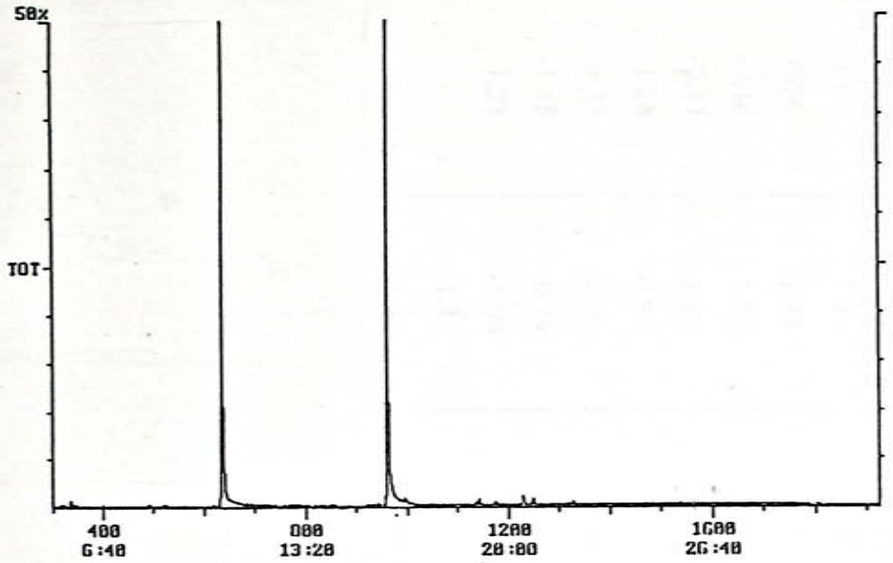
75


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

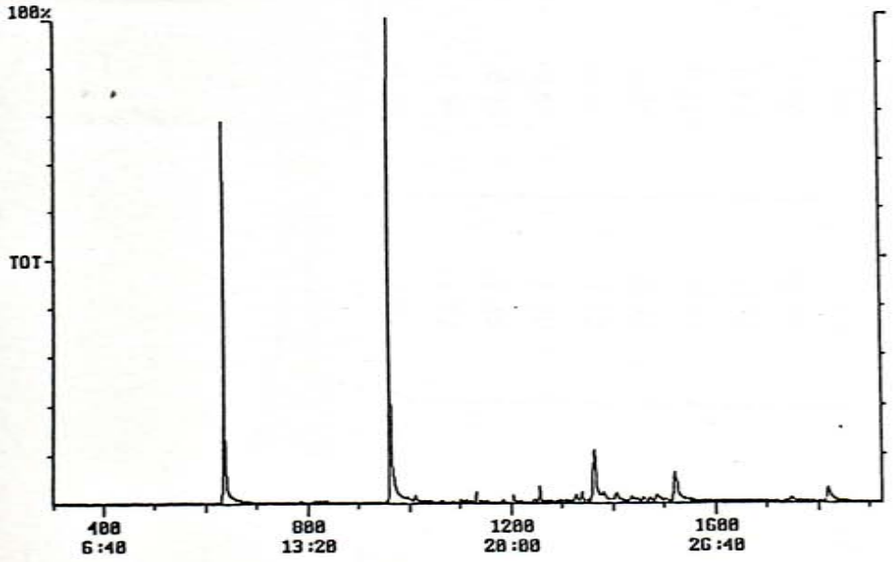
76


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

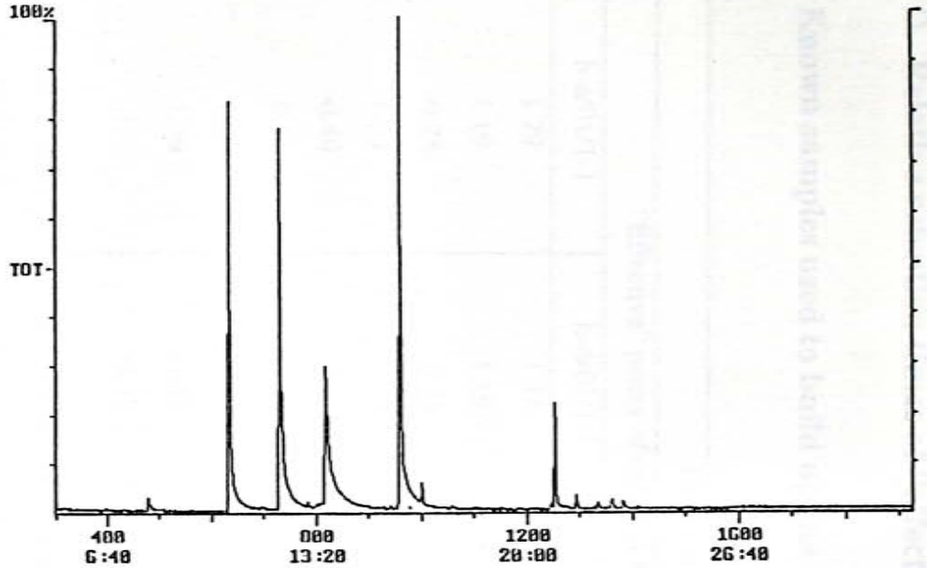
77


50%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

78


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

79


100%
TOT
400
6:40
800
13:20
1200
20:00
1600
26:40

D. Relative retention times of "effective" and "characteristic" peaks

A2f

Known samples used to build up the reference file

Danggui				"characteristic" peaks
"Effective" peaks of samples from different sources				
h-a/01/1-1	h-d/01/1-1	h-e/01/1-1	std/01/1-1	
1.79	1.78	1.78	-0.28	1.78
1.19	1.19	1.19	-0.43	1.19
-0.28	-0.28	-0.28	-0.52	-0.28
-0.31	-0.30	-0.39	-0.56	-0.39
-0.40	-0.39	-0.56	-0.88	-0.56
-0.56	-0.87	-0.87	-1.13	-0.87
-1.33	-1.33	-1.32	-1.33	-1.33
-1.56	-1.56	-1.56	-1.56	-1.56
-	-1.76	-2.02	-	-
-	-	-2.57	-	-

Duohuo					"Characteristic" peaks
"Effective" peaks of samples from different sources					
h-a/02/1-2	h-d/02/1-4	h-e/02/1-2	b-a/02/1-1	std/02/1-1	
1.80	1.90	1.89	1.90	1.48	1.90
0.54	1.82	1.80	1.80	0.54	1.80
-0.46	1.80	-0.30	0.82	-0.36	0.54
-0.87	1.06	-0.46	0.53	-0.87	-0.46
-1.28	0.82	-0.87	-0.36	-0.92	-0.87
-1.35	0.54	-0.96	-0.46	-1.13	-1.35
-1.38	-0.74	-1.19	-0.87	-1.30	-1.53
-1.53	-1.06	-1.35	-0.93	-1.35	-
-	-1.35	-1.53	-1.36	-1.77	-
-	-1.53	-1.77	-1.54	-2.03	-

Qianghuo				"Characteristic" Peaks
"Effective" peaks of samples from different sources				
h-a/04/1-1	b-a/04/1-1	g-a/04/1-1	std/04/1-1	
1.90	1.86	1.90	1.06	1.83
1.83	1.82	1.83	-0.57	1.68
1.81	1.68	1.81	-0.64	1.06
1.68	1.34	1.68	-0.74	0.54
1.55	1.06	1.34	-0.88	-0.56
1.06	0.54	1.06	-0.97	-0.63
0.98	-0.56	0.54	-1.01	-1.22
0.54	-0.63	-0.55	-1.12	-
-0.55	-1.11	-0.62	-1.23	-
-1.21	-1.22	-0.99	-1.62	-

Baizhu				"Characteristic" Peaks
"Effective" peaks of samples from different sources				
h-a/05/1-2	b-a/05/1-1	g-a/05/1-1	std/05/1-1	
-0.13	-0.13	-0.12	-0.68	-0.13
-0.17	-0.17	-0.16	-0.73	-0.17
-0.30	-0.46	-0.46	-0.78	-0.46
-0.46	-0.68	-0.68	-1.19	-0.68
-0.60	-0.74	-0.73	-1.24	-0.73
-0.68	-0.79	-0.78	-1.27	-0.78
-0.72	-0.82	-0.82	-1.47	-0.82
-0.78	-1.24	-1.24	-1.56	-1.24
-0.82	-	-1.55	-2.25	-
-1.24	-	-2.25	-2.73	-

Canzhu					"Characteristic" peaks
"Effective" peaks of samples from different sources					
h-b/06/1-1	h-d/06/1-1	g-a/06/1-1	std/06/1-1	std/06-b/1-1	
0.24	1.89	1.91	0.03	-0.68	0.06
0.06	0.06	0.06	-0.68	-0.74	0.03
0.03	0.02	0.03	-0.9	-0.78	-0.45
-0.08	-0.09	-0.39	-1.11	-1.12	-0.67
-0.13	-0.38	-0.46	-1.15	-1.15	-0.78
-0.45	-0.45	-0.74	-1.22	-1.22	-1.22
-0.67	-0.48	-0.78	-1.24	-1.24	-
-0.77	-0.67	-1.11	-1.27	-1.27	-
-1.23	-0.71	-1.15	-1.46	-1.46	-
-1.55	-1.21	-1.22	-1.55	-1.74	-

Jingjie					"Characteristic" Peaks
"Effective" peaks of samples from different sources					
h-a/07/1-3	h-d/07/1-1	b-a/07/1-2	g-a/07/1-2	std/07/1-1	
1.81	1.81	1.19	1.18	1.80	1.81
1.28	1.19	1.14	1.13	1.19	1.19
1.19	1.14	0.86	1.07	1.14	1.14
1.14	0.76	0.75	0.86	0.88	0.86
1.08	-	-0.87	0.75	0.86	0.76
0.86	-	-1.37	0.68	0.76	-0.90
0.76	-	-1.78	-0.90	-0.88	-
-0.91	-	-2.03	-1.30	-0.90	-
-	-	-2.45	-2.57	-1.77	-
-	-	-2.58	-	-2.02	-

Bajiaohuixiang		"Characteristic" Peaks
"Effective" peaks of samples from different sources		
f-b/08/1-1	std/08/1-1	
1.81	1.82	1.82
0.97	1.79	0.98
0.51	1.73	0.51
-0.12	1.06	-0.17
-0.18	0.98	-
-	0.50	-
-	-0.17	-
-	-0.51	-
-	-0.87	-
-	-1.21	-

Sharen				"Characteristic" Peaks
"Effective" peaks of samples from different sources				
h-a/09/2-3	h-e/09/1-1	g-a/09/1-1	std/09/1-1	
1.81	1.98	1.81	1.22	1.81
1.22	1.81	1.22	1.14	1.22
1.09	1.23	1.09	1.10	1.10
0.54	1.10	0.54	0.54	0.54
-0.88	0.54	0.09	-0.60	-
-	0.10	-0.12	-0.88	-
-	-	-0.49	-0.90	-
-	-	-	-1.03	-
-	-	-	-1.13	-
-	-	-	-1.84	-

Ezhu					"Characteristic" Peaks
"Effective" peaks of samples from different sources					
h-b/11/1-1	h-d/11/1-1	h-e/11-1-1	b-c/11-b/1-1	std/11/1-1	
1.78	1.22	0.03	1.78	1.78	1.22
1.22	0.02	-0.29	1.22	-1.06	0.02
1.14	-0.49	-0.48	1.14	-1.14	-0.49
0.02	-0.99	-0.98	0.02	-1.22	-0.99
-0.49	-1.13	-1.02	-0.42	-1.25	-1.05
-0.78	-1.23	-1.05	-0.49	-1.30	-1.22
-0.99	-1.25	-1.11	-0.99	-1.33	-1.39
-1.04	-1.34	-1.21	-1.22	-1.38	-1.47
-1.07	-1.40	-1.38	-1.39	-1.47	-
-1.39	-1.48	-1.46	-1.48	-1.62	-

Yujin						"Characteristic" Peaks
"Effective" peaks of samples from different sources						
h-d/12/1-1	b-a/12/1-1	g-a/12/1-1	b-c/12-c/1-1	f-b/12-b/1-1	std/12-c/1-1	
0.02	1.22	1.78	1.79	1.22	1.79	1.22
-0.93	1.14	1.22	1.22	0.02	1.22	1.14
-0.99	0.02	1.14	1.14	-0.13	1.14	0.02
-1.08	-0.87	0.02	0.03	-0.42	-1.05	-1.06
-1.13	-1.40	-0.46	-0.48	-0.49	-1.21	-1.39
-1.23	-1.58	-0.49	-1.06	-0.78	-1.28	-
-1.30	-1.61	-0.79	-1.22	-0.87	-1.32	-
-1.34	-1.84	-1.30	-1.38	-0.99	-1.38	-
-2.16	-2.02	-1.39	-1.47	-1.07	-1.46	-
-	-2.57	-1.48	-1.78	-1.39	-1.61	-

Chuanxiong					"Characteristic" Peaks
"Effective" peaks of samples from different sources					
h-a/15/1-2	h-d/15/1-2	h-e/15/1-1	h-f/15/1-1	std/15/1-1	
1.86	1.82	1.83	1.82	1.68	1.82
1.82	1.68	1.68	1.68	1.06	1.68
1.68	1.06	1.55	1.06	-0.16	1.06
1.06	-0.46	1.20	0.98	-0.46	-0.46
-0.46	-0.88	1.06	-0.46	-0.49	-0.88
-1.24	-0.99	-0.47	-1.24	-0.88	-1.24
-1.32	-1.24	-0.88	-1.32	-1.24	-1.32
-1.54	-1.33	-1.24	-1.36	-1.32	-1.53
-1.56	-1.58	-1.32	-1.53	-1.53	-1.56
-	-	-1.56	-1.56	-1.56	-

Qianhu			"Characteristic" Peaks
"Effective" peaks of samples from different sources			
h-b/18/1-1	h-d/18/1-1	std/18/1-1	
1.90	1.81	1.48	1.82
1.82	1.10	1.23	1.06
1.36	1.06	0.90	0.98
1.33	0.98	0.54	-0.10
1.09	0.46	-0.10	-0.56
1.06	-0.06	-0.28	-
0.98	-0.10	-0.55	-
0.84	-0.28	-0.88	-
-0.10	-0.31	-0.90	-
-0.56	-0.56	-1.02	-

Fangfeng			"Characteristic" Peaks
"Effective" peaks of samples from different sources			
h-a/19/1-2	h-d/19/1-2	h-e/19/1-1	
1.98	1.98	1.91	1.98
1.92	1.92	0.49	1.92
1.22	0.54	-0.48	0.54
1.10	-0.28	-0.87	-0.56
1.03	-0.54	-1.36	-0.82
0.54	-0.56	-1.77	-0.90
-0.56	-0.82	-1.99	-
-0.82	-0.90	-2.02	-
-0.90	-2.82	-2.16	-
-	-	-2.58	-

Muxiang					"Characteristic" Peaks
"Effective" peaks of samples from different sources					
h-d/20/1-3	h-e/20/1-1	f-a/20/1-2	g-a/20/1-2	std/20/1-1	
0.54	-0.11	-0.15	-0.11	-0.74	-0.11
-0.06	-0.15	-0.90	-0.15	-0.90	-0.15
-0.11	-0.42	-1.24	-0.90	-1.12	-0.90
-0.15	-0.46	-1.30	-1.24	-1.16	-1.24
-0.49	-1.12	-0.36	-1.71	-1.24	-1.36
-1.24	-1.24	-1.61	-2.25	-1.29	-1.72
-1.36	-1.36	-1.71	-2.38	-1.37	-2.41
-1.72	-1.72	-2.38	-2.58	-1.46	-2.63
-2.41	-2.26	-2.41	-	-1.73	-
-2.63	-2.41	-2.63	-	-2.63	-

Zisuye (purplish green)				"Characteristic" Peaks
"Effective" peaks of samples from different sources				
h-a/23/1-5	h-d/23/1-4	h-e/23/1-1	h-f/23/1-1	
0.70	0.70	1.07	-0.12	-0.13
-0.13	-0.13	-0.13	-0.27	-0.27
-0.28	-0.27	-0.27	-0.30	-0.30
-0.30	-0.30	-0.30	-0.49	-0.90
-0.91	-0.90	-0.42	-0.90	-1.05
-1.04	-1.05	-0.49	-1.05	-1.31
-1.32	-1.15	-0.90	-1.31	-
-1.37	-1.31	-1.05	-1.74	-
-1.78	-	-1.31	-	-
-	-	-1.74	-	-

Zisuye (greenish)			"Characteristic" Peaks
"Effective" peaks of samples from different sources			
b-a/23/1-2	g-a/23/1-1	std/23/1-1	
1.47	1.48	1.47	1.47
0.82	0.70	0.70	0.70
0.70	-0.13	0.43	0.44
0.49	-0.30	-0.12	-0.13
0.44	-0.88	-0.88	-0.88
-0.13	-0.90	-0.90	-0.90
-0.88	-1.08	-1.03	-1.23
-0.90	-1.24	-1.15	-1.30
-1.30	-1.30	-1.23	-
-2.56	-1.32	-1.30	-

Xiangru			"Characteristic" Peaks
"Effective" peaks of samples from different sources			
h-a/24/1-1	h-d/24/1-1	b-a/24/1-1	
1.83	1.83	1.83	1.83
1.79	1.06	1.80	1.80
0.48	0.48	1.23	0.48
0.43	0.43	1.06	0.43
0.23	0.23	0.48	0.23
-0.12	0.13	0.43	0.13
-0.18	-0.12	0.23	-0.12
-0.30	-0.18	0.13	-0.30
-1.02	-0.30	-0.30	-1.03
-	-1.03	-1.03	-

Gaoliangjiang			"Characteristic" Peaks
"Effective" peaks of samples from different sources			
h-a/25/1-1	h-d/25/1-1	b-b/25/1-1	
1.81	1.82	2.02	1.81
1.78	1.80	1.81	1.79
1.22	1.23	1.79	1.22
1.06	1.10	1.22	1.06
0.98	1.06	1.06	0.98
-0.18	0.98	0.98	-0.18
-0.49	-0.04	0.88	-0.49
-0.56	-0.17	-0.18	-0.56
-0.59	-0.49	-0.49	-
-0.70	-0.56	-0.56	-

Peilan			"Characteristic" Peaks
"Effective" peaks of samples from different sources			
h-a/29/1-1	b-b/29/1-1	f-a/29/1-1	
0.82	0.49	0.46	-0.10
-0.10	0.41	-0.10	-0.13
-0.12	0.17	-0.13	-0.39
-0.18	0.03	-0.40	-0.43
-0.39	-0.13	-0.43	-0.59
-0.43	-0.30	-0.62	-0.61
-0.59	-0.42	-0.88	-0.90
-0.61	-0.59	-0.90	-0.97
-0.90	-0.87	-0.97	-
-0.97	-0.90	-1.30	-

Huoxiang			"Characteristic" Peaks
"Effective" peaks of samples from different sources			
h-a/30/1-1	b-a/30/1-1	f-b/30/1-1	
0.05	0.05	0.05	0.05
-0.13	-0.19	-0.20	-0.19
-0.19	-0.27	-0.27	-0.27
-0.27	-0.32	-0.33	-0.33
-0.33	-0.36	-0.52	-0.36
-0.36	-0.51	-0.86	-0.51
-0.51	-0.85	-0.90	-0.85
-0.85	-1.09	-1.10	-1.09
-1.23	-1.23	-1.23	-1.23
-1.29	-1.29	-1.03	-1.29